Atmospheric and Environmental Research, Inc.

Algorithm Theoretical Basis Document (ATBD) for the Conical-Scanning Microwave Imager/Sounder (CMIS) Environmental Data Records (EDRs)

Volume 3: Water Vapor EDRs

Covering: Atmospheric Vertical Moisture Profile EDR
Precipitable Water EDR

Version 1.2 – 15 March 2001

Solicitation No. F04701-01-R-0500

Submitted by:
Atmospheric and Environmental Research, Inc.
131 Hartwell Avenue
Lexington, MA 02421-3126

With contributions by:
Alan Lipton, Jean-Luc Moncet, Jennifer Hegarty, Sid-Ahmed Boukabara, Hélène Rieu-Isaacs, Xu Liu, Richard Lynch, John Galantowicz

Prepared for:
Boeing Satellite Systems
919 Imperial Avenue
El Segundo, CA 90245

This page intentionally left blank.

REVISION HISTORY

Version	Release Date	POC	Comments
1.0		J. L. Moncet, AER, Inc.	First Draft Version
1.1			Revised requirements section Rewrote performance section
1.2	1/21/01		Added performance material for cirrus cloud conditions and regarding surface emissivity. Revised performance under cloudy and degraded measurement conditions

RELATED CMIS DOCUMENTATION

Government Documents

Title	Version	Authorship	Date
CMIS SRD for NPOESS Spacecraft and Sensors	3.0	Associate Directorate for Acquisition NPOESS IPO	2 March 2001

Boeing Satellite Systems Documents

Title		Covering
ATBD for the CMIS TDR/SDR Algorithms		
ATBD for the CMIS EDRs	Volume 1: Overview	Part 1: Integration
		Part 2: Spatial Data Processing • Footprint Matching and Interpolation • Gridding • Imagery EDR
	Volume 2: Core Physical Inversion Module	
	Volume 3: Water Vapor EDRs	Atmospheric Vertical Moisture Profile EDR Precipitable Water EDR
	Volume 4: Atmospheric Vertical Temperature Profile EDR	
	Volume 5: Precipitation Type and Rate EDR	
	Volume 6: Pressure Profile EDR	
	Volume 7: Cloud EDRs	Part 1: Cloud Ice Water Path EDR
		Part 2: Cloud Liquid Water EDR
		Part 3: Cloud Base Height EDR
	Volume 8: Total Water Content EDR	
	Volume 9: Soil Moisture EDR	
	Volume 10: Snow Cover/Depth EDR	

Title		Covering
	Volume 11: Vegetation/Surface Type EDR	
	Volume 12: Ice EDRs	Sea Ice Age and Sea Ice Edge Motion EDR Fresh Water Ice EDR
	Volume 13: Surface Temperature EDRs	Land Surface Temperature EDR Ice Surface Temperature EDR
	Volume 14: Ocean EDR Algorithm Suite	Sea Surface Temperature EDR Sea Surface Wind Speed/Direction EDR Surface Wind Stress EDR
	Volume 15: Test and Validation	All EDRs

Bold = this document

TABLE OF CONTENTS FOR VOLUME 3

LIST OF TABLES

LIST OF FIGURES

TABLE OF FIGURES

TABLE OF FIGURES

1. Introduction

1.1. Purpose

This algorithm theoretical basis document (ATBD) provides the underlying mathematical and theoretical background for the Atmospheric Vertical Moisture Profile (AVMP) and Precipitable Water (PW) EDRs (Environmental Data Records) for the Conical-Scanning Microwave Imager/Sounder (CMIS) developed by Atmospheric and Environmental Research, Inc. (AER) in support of the National Polar-orbiting Operational Environmental Satellite System (NPOESS).

The water vapor vertical profile and precipitable water EDR requirements support a number of diverse applications including specification of atmospheric transmission and propagation for a variety of sensor and communications systems, specification of relevant initialization fields for meso- and global scale numerical weather, hydrological, and cloud forecasting models, and studies of atmospheric transport of water vapor. As a key component of the atmospheric water substance budget (including vapor, liquid, ice, and various forms of precipitation), specification of these EDRs is also a factor in the overall hydrological cycle. Water vapor plays a key role in radiation transport and is, thus, also significant as a climate system variable.

The extraction of water vapor profiles from the CMIS measurements are the driving requirements for our algorithm design. Our approach (cf ATBD for CMIS EDRs – Volume 1: Overview - Part 1: Integration and CMIS EDR Algorithm ATBD – Volume 2: Core Physical Inversion Module) is best suited for producing water vapor retrievals in clear and cloudy conditions (both liquid water and cirrus ice) and over all surface types, therefore maximizing the range of conditions over which the related EDRs as Atmospheric Vertical Moisture Profile and Precipitable Water will be made available. The objective of this ATBD is to facilitate an understanding of our approach to the CMIS retrieval problem from a phenomenological perspective in the context of the current state-of-the-art. The relationship between the Core Module and the AVMP and PW EDR algorithms is detailed in the ATBD for CMIS EDRs, Volume 1: Overview, Part 1: Integration.

1.2. Scope

Section 2 explains the physics of the problem and describes the relevant CMIS characteristics involved in the process. Section 3 describes the core physical inversion algorithm processing flow with its input and output data. Section 4 presents the theoretical and mathematical description of the algorithm. The algorithm performance is given in section 5 with a description of the test data followed by sensitivity studies, performance of the algorithm under different conditions, and a description of the constraints and limitations of the algorithm. Four appendices follow on the sounding Channel optimization (Appendix 1), the definition of the 183 GHz channel set (Appendix 2), an analysis of the 183 GHz channel set (Appendix 3) and the sensitivity of 183 GHz channels to water vapor (Appendix 4). References and a list of acronyms follow.

2. Overview and Background Information

2.1. Objectives of the Atmospheric Vertical Moisture Profile and Precipitable Water EDR Retrieval

The Atmospheric Vertical Moisture Profile (AVMP) algorithm must produce estimates of the water vapor of the atmosphere. The water vapor reports are to be made along vertical paths through the atmosphere, using the term profile to refer to a set of water vapor mixing ratios along a single path. Each report consists of water vapor mixing ratios given as a function of atmospheric pressure for a specified location. Water vapor profiles are to be produced within the swath observed by CMIS so that coverage is global upon a series of NPOESS orbits.

For a conical scanner such as CMIS, the paths over which the instrument views the atmosphere are slanted with respect to the local vertical of the observed location. The AVMP algorithm produces estimates of water vapor along the CMIS view paths, given as a function of pressure, as an initial algorithm product. These water vapor profiles then undergo further processing to generate the final EDR products that are registered to vertically-oriented paths, for compliance with the CMIS System Requirements Document (SRD).

The Precipitable Water (PW) algorithm produces an estimate of the total equivalent water in a vertical column of the atmosphere per unit cross-sectional.

2.2. Summary of AVMP and PW EDRs Requirements

2.2.1. Atmospheric Vertical Moisture Profile

The text below and Table 2-1 are the portions of CMIS SRD section 3.2.1.1.1.1 that apply directly to the AVMP algorithm.

An atmospheric vertical moisture profile is a set of estimates of the average mixing ratio in three-dimensional cells centered on specified points along a local vertical. The mixing ratio of a sample of air is the ratio of the mass of water vapor in the sample to the mass of dry air in the sample. For requirements in which both a percentage value and an absolute value are supplied in the table below, the requirement is to be interpreted as the greater of the values.

Table 2-1: Atmospheric Vertical Moisture Profile EDR Requirements.

Para. No.	Description	Thresholds	Objectives
C40.2.1-1	a. Horizontal Cell Size	15 km	2 km
C40.2.1-2	b. Horizontal Reporting Interval	15 km	2 km
C40.2.1-3	c. Vertical Cell Size	2 km	2 km
	d. Vertical Reporting Interval		
C40.2.1-4	1. Surface to 850 mb	20 mb	5 mb
C40.2.1-5	2. 850 mb to 100 mb	50 mb	15 mb
C40.2.1-6	e. Horizontal Coverage	Global	Global
C40.2.1-7	f. Vertical Coverage	Surface to 100 mb	Surface to 100 mb
C40.2.1-8	g. Measurement Range	0 - 30 g/kg	0 - 30 g/kg
	h. Measurement Uncertainty (expressed as a percent of average mixing ratio in 2 km layers)		
	Clear		
C40.2.1-9	1. Surface to 600 mb	20 % or 0.2 g/kg	10 %
C40.2.1-10	2. 600 mb to 300 mb	35 % or 0.1 g/kg	10 %
C40.2.1-11	3. 300 mb to 100 mb	35 % or 0.04 g/kg	10 %
	Cloudy		
C40.2.1-12	4. Surface to 600 mb	20 % or 0.2 g/kg	10 %
C40.2.1-13	5. 600 mb to 300 mb	40 % or 0.1 g/kg	10 %
C40.2.1-14	6. 300 mb to 100 mb	40 % or 0.04 g/kg	10 %
C40.2.1-15	i. Mapping Uncertainty	5 km	1 km
C40.2.1-16	j. Swath Width	1700 km	3000 km

2.2.2. Precipitable Water

The text below and Table 2-2 are the portions of CMIS SRD section 3.2.1.1.1.1 that apply directly to the Precipitable Water algorithm.

The requirements in Table 2-2 apply under both clear and cloudy conditions. Precipitable water is defined as the total equivalent water in a vertical column of the atmosphere per unit cross-sectional area.

Table 2-2: Precipitable Water EDR Requirements.

Para. No.	Description	Thresholds	Objectives
C40.3.3-1	a. Horizontal Cell Size	25 km (TBR)	1 km
C40.3.3-2	b. Horizontal Reporting Interval	25 km (TBR)	1 km
C40.3.3-3	c. Horizontal Coverage	Global	Global
C40.3.3-4	d. Measurement Range	0 - 75 mm	0 - 100 mm
C40.3.3-5	e. Measurement Accuracy	Greater of 10 % or 2 mm	1 mm
C40.3.3-6	f. Measurement Precision	1 mm	1 mm
C40.3.3-7	g. Mapping Uncertainty	3 km	0.1 km
C40.3.3-8	h. Swath Width	1700 km	3000 km

2.2.3. Interpretation of SRD Requirements

Revisions to the SRD have clarified that the horizontal cell size requirement is meant to specify the area over which averaging is done when validating EDRs against truth data. We infer that the AVMP vertical cell size requirement is, likewise, a validation requirement. According to our interpretation, the vertical cell size represents the distance over which profile data are averaged when validating EDRs against truth data. We infer that the EDR product to be delivered to the CMIS customers should be a profile of point values with the greatest vertical resolution the CMIS system is capable of providing (within the context of other system requirements), without any deliberate vertical averaging (smoothing). Only in the course of validation are the AVMP data processed so that they are estimates of the average mixing ratio in three-dimensional cells.

The PW measurement error is specified in terms of accuracy and precision. We interpret these, in accordance with the SRD glossary (Appendix A) as the bias and standard deviation of the measurement error, respectively, for data that have been binned according to the true value of PW. We believe that the numeric values that were applied to these two metrics should be exchanged in order to be consistent with the definitions of the metrics and the error characteristics of environmental remote sensing products.

2.3. Physics of the Problem

The sensitivity of microwave radiometric measurements, fundamentally, arises from the fact that water vapor absorbs microwave radiation and emits radiation in relation to the air temperature. Some insight in the retrieval problem can be gained by considering a simplified version of the

radiative transfer equation treating the problem as monochromatic, non-scattering, with the Rayleigh-Jeans approximation, and neglecting reflected radiation:

$$T_B = \varepsilon_s T_s \Im_s + \int_{\tau_s}^{1} T \, d\Im \tag{1}$$

where T_B is the brightness temperature, T is the surface temperature, $\Im$ is the transmittance to space, *s* denotes the surface, and ε_s is the surface emissivity. If we break out a discrete layer just above the surface, the equation becomes:

$$T_B = \varepsilon_s T_s \Im_s + \int_{\tau_s}^{\tau_{TOP}} T \, d\Im + \int_{\tau_{TOP}}^{1} T \, d\Im \tag{2}$$

and, with finite differencing,:

$$T_B = \varepsilon_s T_s \Im_s + T_l \left(\Im_{TOP} - \Im_s \right) + T_a \left(1 - \Im_{TOP} \right) \tag{3}$$

where *TOP* refers to the top of the layer and T_l and T_a refer to the effective temperature of the layer and the atmosphere above it, respectively. This can be rewritten as:

$$T_B = \varepsilon_s T_s e^{-(\tau_l + \tau_a)} + T_l \left(e^{-\tau_a} - e^{-(\tau_l + \tau_a)} \right) + T_a \left(1 - e^{-\tau_a} \right) \tag{4}$$

where τ is the optical depth of a layer. If we take the derivative of the brightness temperature with respect to the optical depth, we obtain:

$$\frac{\partial T_B}{\partial \tau_l} = -\varepsilon_s T_s e^{-(\tau_l + \tau_a)} + T_l \left(e^{-(\tau_l + \tau_a)} \right) = \left(T_l - \varepsilon_s T_s \right) e^{-(\tau_l + \tau_a)} = \left(T_l - \varepsilon_s T_s \right) \Im_s \tag{5}$$

The sensitivity to water vapor is represented by the derivative with respect to water vapor amount:

$$\frac{\partial T_B}{\partial q_l} = \frac{\partial T_B}{\partial \tau_l} \frac{\partial \tau_l}{\partial q_l} = \left(T_l - \varepsilon_s T_s \right) \Im_s \frac{\partial \tau_l}{\partial q_l} \tag{6}$$

The water vapor signal approaches zero whenever any of the terms on the right approach zero. The first term on the right is the contrast between the temperature of the layer and the effective temperature of the surface background. This term may be positive or negative, and may be small when $T_l \approx T_s$ and $\varepsilon_s \approx 1$, as is often the case over land surfaces. The term $\Im_s$ is the transparency of the atmosphere above the layer of interest and $\frac{\partial \tau_l}{\partial q_l}$ is the strength of the water vapor absorption. For the PW EDR, most of the relevant water vapor is in a layer just above the

surface in most cases. For water vapor profiling, we would get a similar set of equations if we were to break out a layer at a higher altitude, but the contrast term would compare the temperature of the layer and the temperature of the atmosphere below the layer. The sensitivity thus goes to zero in isothermal layers that are deep in relation to the vertical resolution of the water vapor sounding channel.

There are a variety of factors that can interfere with detection of any water vapor signal that may be present. The factors can be highlighted by taking the equation for the brightness temperature and considering the variation in the brightness temperature that results from variations in the terms

$$\delta T_B = \frac{\partial T_B}{\partial \tau_l}\delta\tau_l + \frac{\partial T_B}{\partial \tau_a}\delta\tau_a + \frac{\partial T_B}{\partial \varepsilon_s}\delta\varepsilon_s + \frac{\partial T_B}{\partial T_s}\delta T_s + \frac{\partial T_B}{\partial T_l}\delta T_l + \frac{\partial T_B}{\partial T_a}\delta T_a \qquad (7)$$

The terms on the right side are identified in Table 2-3, including the method to mitigate the effect with the channel set. Scattering by precipitation, which was not included in the equation, may be an additional large source of interference.

Table 2-3. Factors affecting variation of brightness temperatures

	Factor	*Mitigation with channel set*
1	Uncertainty of cloud or other absorbers	Include channels where water vapor signal is stronger than cloud signal
2	Uncertainty of water vapor or cloud in other layers	Include channels with sensitivity concentrated in other layers
3	Emissivity uncertainty	Include nearby window channels, with similar emissivity and less water vapor sensitivity Use emissivity spectrum constraint
4	Surface temperature uncertainty	
5	Uncertainty of temperature in layer	Include temperature sounding channels
6	Uncertainty of temperature in other layers	

More discussion is given in the Core Module ATBD on the physics of the problem and how can water vapor be retrieved from the CMIS instrument.

2.4. Instrument Characteristics

The primary CMIS channels for AVMP and PW are one channel along the weak 23 GHz water vapor line (optimized within NTIA allocations) and three channels along the 183 GHz water vapor line (optimized for water vapor sounding skill – see Appendicies 1, 2, and 3 on Analysis of 183 GHz Channel Set). The secondary channels are the window channels at 18, 36, 89, and 166 GHz and the temperature sounding channels at 50-60 GHz.

The 183 GHz measurements provide information on water vapor concentration at different depths in the troposphere, given the knowledge of atmospheric temperature profile. These channels are also sensitive to effects of cloud liquid water and cloud ice (for large particles). The weak 22.235 GHz is used for lower tropospheric and total column water vapor.

In the lower troposphere, specification of surface emissivity and skin temperature is also required. Requirements for measurement uncertainty can only be achieved with an algorithm that adequately distinguishe between molecular water and liquid/ice water. Over open ocean, where surface emissivity is relatively stable, an emissivity constraint can be derived from the emissivity in the lower frequency channels or by direct use of wind speed/direction and SST information derived from CMIS. Over land, emissivity is more highly variable and such correlation may not exist, for certain surface types. Therefore, the CMIS algorithm/sensor system should have the capability of correcting for the surface effects at the high frequencies.

The main channels for precipitable water retrieval are located on the 22.235 line. The 22 GHz measurements only have a second order dependence on vertical distribution of water vapor. Therefore, the profile information obtained from the moisture sounding channels (183 GHz) has minimal impact on the system performance over ocean. The water vapor channels should be located off the line center to avoid problems of line saturation in the tropics and minimizes effects of uncertainties in the line width. The baseline channel selection for precipitable water consists of 23.8 V and 23.8 H channels. Additional information about surface emissivity and skin temperature, atmospheric temperature and cloud liquid water must be provided. These additional unknowns are retrieved together with the precipitable water EDR from the combination of 18.7, 23.8 and 37 GHz channels. Over land, loss of surface/atmosphere contrast degrades the sensitivity to water vapor and impairs ability of system to separate cloud liquid

water from water vapor. Temperature and water vapor sounding channels have a significant positive impact on system performance for precipitable water over land.

2.5. Requirements for Cross Sensor Data

The AVMP and precipitable water EDRs require real-time cross-sensor data from VIIRS for cloud top and cloud cover, with the characteristics listed in the last two rows of Table 2-4. The cloud cover data are used to detect clear conditions over land, eliminating the radiometric ambiguity between cloud and water vapor and improving water vapor retrieval performance. The cloud top data are used to reduce the radiometric ambiguities between cirrus clouds and water vapor. The cloud data are also used in the quality control process (ATBD Vol. 2) to aid in identifying and flagging conditions with precipitation. The AVMP and precipitable water EDRs also require that data from VIIRS and CrIS are used in the development and maintenance of the dynamic emissivity database that is used in the core module. For this purpose, the data need not be from the current orbit, but may be from previous recent orbits.

Table 2-4: Requirements for cross-sensor data

Name	Description	Source	Measurement Uncertainty	Swath Width (km)	Horizontal Cell Size (km)	Mapping Uncertainty (km)	Timeliness
LST	Land Surface Temperature	VIIRS	2.5 K rms on a global basis	1700	15	4	≥5 samples within previous 10 days from any satellite with VIIRS and CMIS
WVPF	Atmospheric Moisture Profile (Surface to 600 mb)	CrIS	15%	1700	25	5	≥5 samples within previous 10 days from any satellite with VIIRS and CMIS
PCldTop	Cloud Top Pressure	VIIRS	120 mb uncertainty	1700	15	4	first 5% available within 5 min, last 5% available within 19 min, with even timing of 5% segments; from same satellite
CldCvr	Cloud Cover	VIIRS	80% probability of detection	1700	15	4	first 5% available within 5 min, last 5% available within 19 min, with even timing of 5% segments; from same satellite

2.6. Requirements for External Data

The only external data required to achieve threshold performance for AVMP and Precipitable Water is surface pressure, derived from combination of NWP model forecast data and terrain heights from a high-spatial-resolution global topography database.

2.7. Derived Requirements on Data from other EDR Algorithms

The AVMP algorithm requires view-path water vapor profile data from the Core Module with the characteristics specified in Table 2-5. The AVMP algorithm also uses view-path temperature profile data from the Core Module, for vertical registration, but places no practical requirement on its measurement uncertainty.

Table 2-5: Water Vapor Data Requirements Placed on the Core Module by the AVMP Algorithm.

Parameter	Requirement
a. Horizontal Spatial Resolution	16 km
b. Horizontal Reporting Interval	12.5 km
c. Vertical Cell Size	2 km
d. Vertical Reporting Interval	
1. Surface to 850 mb	20 mb
2. 850 mb to 100 mb	50 mb
e. Horizontal Coverage	Global
f. Vertical Coverage	Surface to 100 mb
g. Measurement Range	0 - 30 g/kg
h. Measurement Uncertainty (expressed as a percent of average mixing ratio in 2 km layers)	
Clear	
1. Surface to 600 mb	20 % or 0.2 g/kg
2. 600 mb to 300 mb	35 % or 0.1 g/kg
3. 300 mb to 100 mb	35 % or 0.04 g/kg
Cloudy	
4. Surface to 600 mb	20 % or 0.2 g/kg
5. 600 mb to 300 mb	40 % or 0.1 g/kg
6. 300 mb to 100 mb	40 % or 0.04 g/kg
i. Mapping Uncertainty	5 km
j. Swath Width	1700 km

3. Algorithm Description

3.1. Historical and Background Perspective of Proposed Algorithms

The problem of retrieving water vapor profiles from passive radiometry is a highly non-linear one. There is a long history of treating the retrieval of water vapor profiles using non-linear physical inversion methods (e.g. Wilheit [Wilheit, xxx], Stephens [Stephens, xxx], UR (Unified Retrieval, [Isaacs, 1987], [Moncet and Isaacs, 1992], [Moncet, Isaacs and Hegarty, 1996]). Physical inversion approaches are also used operationally in weather forecast centers such as ECMWF for the assimilation of water vapor retrieved from AMSU (Advanced Microwave Sounding Unit) and infrared sounders.

In the proposed baseline CMIS algorithm, the moisture profile EDR is derived from the core physical inversion algorithm (see more details in CMIS EDR Algorithm ATBD – Volume 2: Core Physical Inversion Module). The core physical inversion module is required mainly for producing accurate estimates of water vapor profiles for the Atmospheric Vertical Moisture Profile EDR (AVMP), a high priority EDR for CMIS. The Core Module is also required for providing retrieval capability over land for AVMP and related EDRs (Precipitable Water and Total Water Content). The Core Module is a non-linear physical inversion algorithm designed to simultaneously retrieve temperature and water vapor profiles along the sensor view path, cloud liquid water (CLW), cloud top pressure, cloud thickness, skin temperature and spectral surface emissivity.

Precipitable water can be retrieved explicitly from CMIS data or derived by vertical integration of retrieved water vapor mixing ratios. When the same channels are used in the inversion, the two approaches were found to be essentially equivalent. The former has traditionally been the preferred approach for SSM/I retrievals (for which sounding channels are not available) over ocean [Alishouse, 1990; Wentz, xxx] or land [Grody, xxx], independently of whether a statistical or a physical approach was used for inverting the data. In the current CMIS algorithm, because water vapor profile retrievals are required at each measurement location, the latter approach is preferred. It does not require handling precipitable water as a separate parameter and/or the need for a separate inversion. In addition, consistency between the AVMP and PW EDRs both at the inversion and Q/C level is automatically maintained. Note that PW retrievals are still possible with this method even when the higher frequency channels are unavailable. However, in this

case, the number of eigenvectors used in the representation of the vertical mixing ratio profiles needs to be reduced accordingly to ensure the stability of the brightness temperature inversion.

3.2. Theoretical Description

3.2.1. Atmospheric Vertical Moisture Profile EDR Algorithm

Slant-path water vapor profiles are calculated by the core physical inversion module (see more details in the Core Module ATBD).

A second component of the algorithm (vertical re-mapping) takes a set of these slant-path moisture profiles and performs an interpolation process to register the profile data into alignment with the local vertical. The vertical registration process is illustrated in Figure 3-1. The details of the interpolation algorithm are given in the Overview ATBD; Part 2: Footprint Matching and Interpolation.

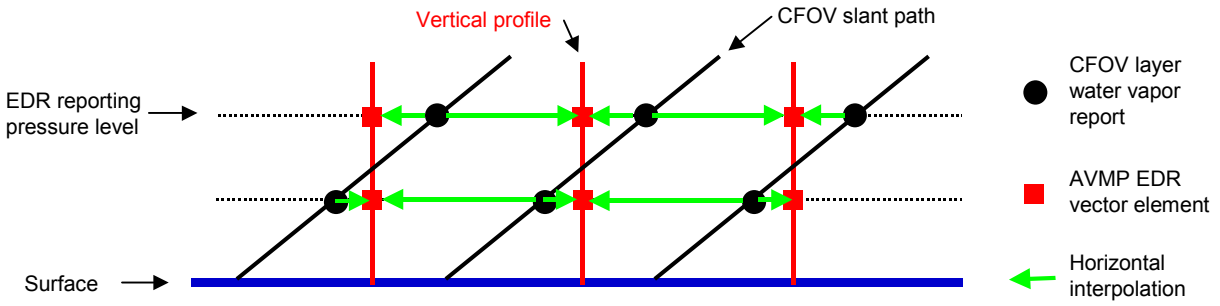


Figure 3-1: Illustration of the process of vertical registration of the water vapor profile, for a cross-sectional view through a portion of a scan.

The profile retrieval step is performed separately and before the vertical registration step in order for the retrievals to have the greatest possible fidelity to the radiometric data, avoiding the introduction of interpolation errors. The slant-path retrievals are then available to be disseminated to CMIS customers that place paramount importance on minimizing water vapor profile errors, and for whom slant-path data are fully acceptable. Such customers may, for example, interpolate the slant-path profiles obtained from CMIS directly to their own grid. Customers that require vertically-registered data may access the final EDR products.

3.2.2. Precipitable Water EDR algorithm

Precipitable water is calculated by vertical integration of Core Module water vapor. It assumes that the logarithm of specific humidity varies linearly with respect to the logarithm of pressure between Core Module pressure levels, which is more consistent with natural trends than the usual assumption that the humidity is constant over the layer (the mean value from the layer top and bottom).

First, specific humidity w is computed from mixing ratio q as:

$$w_i = \frac{q_i}{1+q_i} \tag{8}$$

Then the layer water vapor is computed from the equation set:

$$\begin{cases} \eta_i = p_{i+1} w_{i+1} \ln\left(\dfrac{p_{i+1}}{p_i}\right) \\ \\ \zeta_i = \dfrac{p_i w_i}{p_{i+1} w_{i+1}} \end{cases} \tag{9}$$

so that:

$$\begin{cases} \sigma_i = \dfrac{\eta_i(\zeta_i - 1)}{\ln(\zeta_i)} \quad \text{for} \quad |\zeta_i - 1| > 1\times10^{-5} \\ \\ \sigma_i = \dfrac{2\eta_i}{3-\zeta_i} \quad \text{otherwise} \end{cases} \tag{10}$$

The precipitable water is then computed by summing over all layers:

$$PW = \frac{1}{g}\sum_{i=1}^{N-1}\sigma_i \tag{11}$$

where p is the pressure and g the gravitational acceleration.

3.3. Algorithm Processing Flow

3.3.1. Processing Flow for AVMP Algorithm

The processing flow for the Core Module is illustrated in the *ATBD for the Core Physical Inversion Module* (AER, 2000). The process for vertical registration of the profile data is illustrated in Figure 3-2.

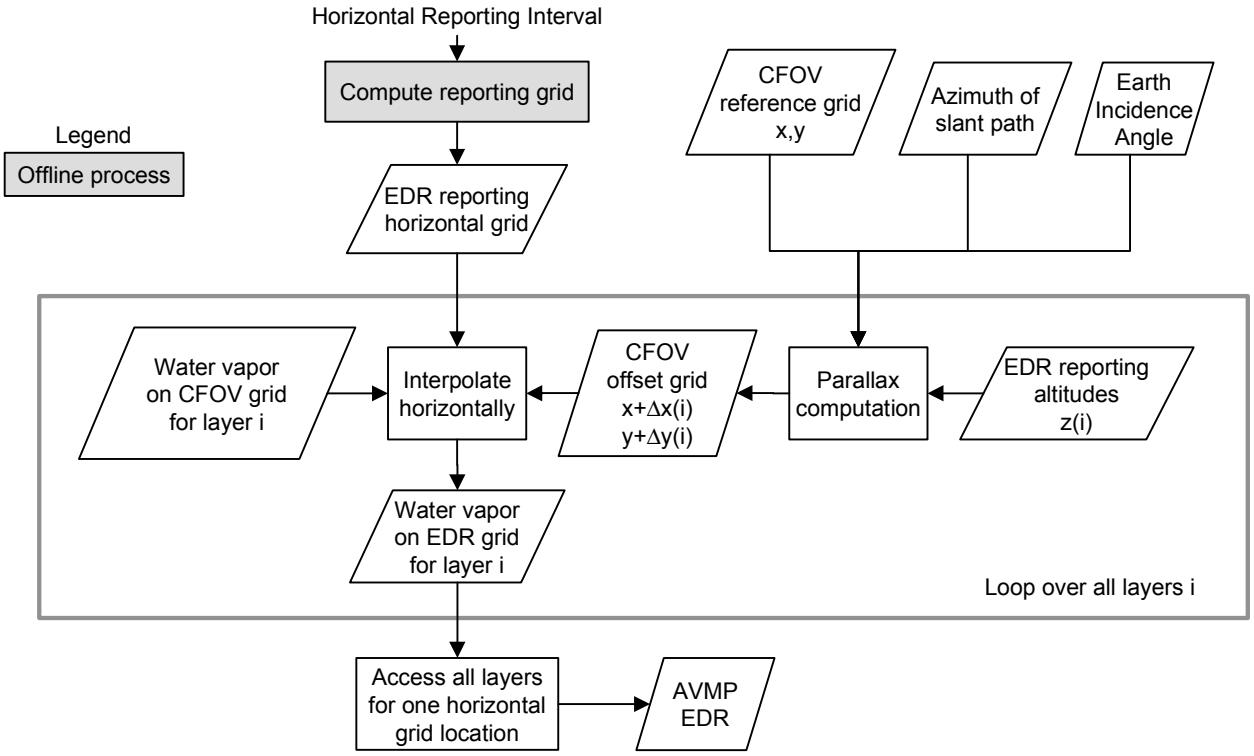


Figure 3-2. Processing flow for vertical registration for the AVMP EDR algorithm.

3.3.2. Processing Flow for PW Algorithm

The process for converting the output from the core module into the PW EDR is illustrated in Figure 3-3.

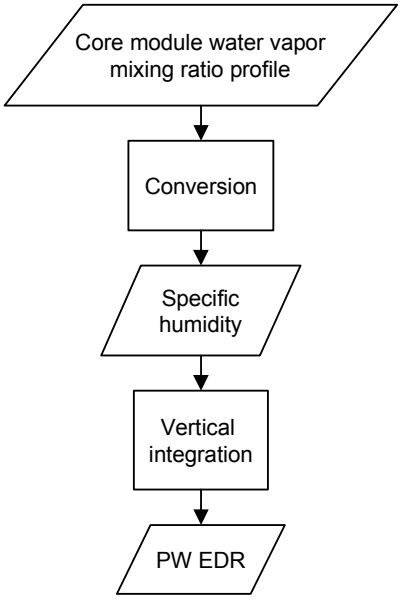


Figure 3-3. Processing flow for the Precipitable Water EDR.

3.3.3. Integration of AVMP and PW Algorithms in the CMIS Algorithm

Figure 3-4 represents a branch of the integrated CMIS EDR Algorithm (more details about the EDR algorithms integration are given in the ATBD for CMIS EDRs – Volume 1: Overview - Part 1: Integration). It is to be read from left to right (time line). The cyan boxes represent the Core Physical Retrieval. Purple shadowed boxes refer to algorithm modules that operate on multiple cells (e.g. vertical re-mapping of sounding product). The AVMP and PW EDR products are on the right.

The required horizontal cell size for the Atmospheric Vertical Moisture Profile EDR is 15 km. As shown in Figure 3-4, the EDR algorithm does not call for water vapor retrievals to be performed directly at 15 km. The design calls for performing retrievals first at 40 or 50-km cells and then using those results (interpolated to 15-km) as a constraint on retrievals at 15-km cells. Before commencing a 15-km retrieval, the brightness temperatures computed from the interpolated data are compared with the observed data. If the brightness temperatures differences are small enough, the 15-km retrieval does no work, but just passes on the water vapor profile that was retrieved at 40 or 50 km. The difference will generally be small when the horizontal gradients of water vapor (and other variables) are small. The difference will be larger, and the 15-km algorithm will be executed, when the gradients are relatively large. The 15-km algorithm

is executed only when the brightness temperature differences indicate that there is more to be gained, with respect to resolving gradients, than there is to be lost, with respect to increased noise in the retrieval. Retrieval performance in areas with small horizontal gradients will be similar to the performance shown for 50-km cells. In areas with larger horizontal gradients, performance will approach that shown for 15-km cells.

The cascade architecture (indicated in Figure 3-4 by the thick arrows connecting the Core Modules operating on different cell sizes) provides a mechanism for incorporating the information from the lower frequency channels in the higher resolution retrievals, and for improving AVMP performance in relatively homogeneous conditions over what would be achievable using a direct inversion of the 15 km radiances.

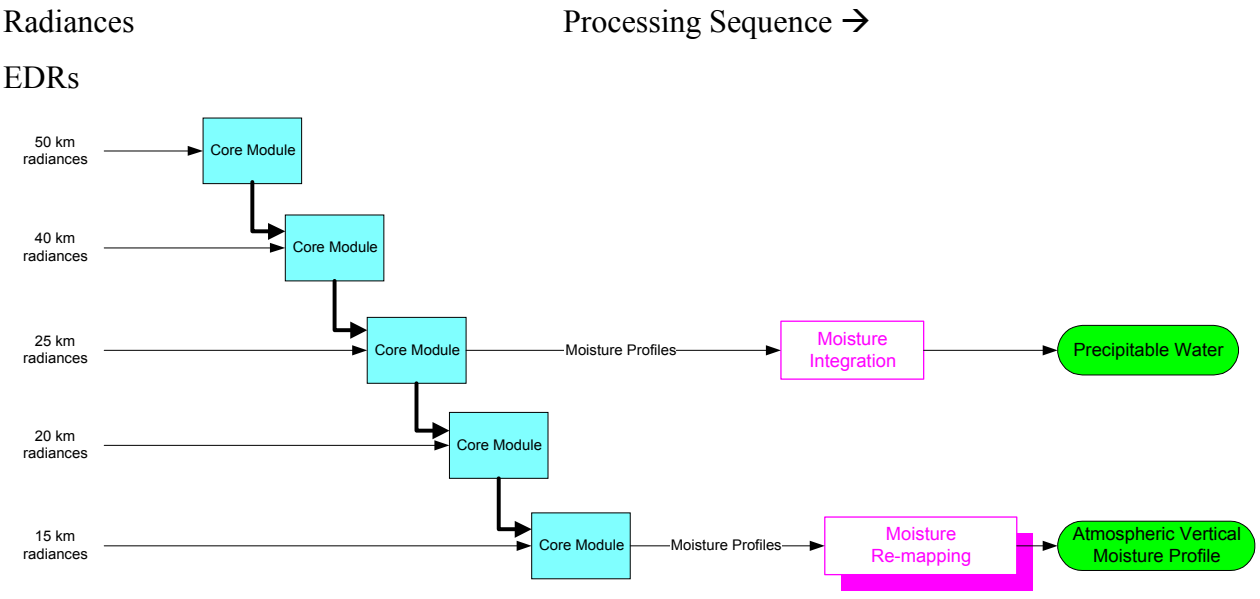


Figure 3-4: Integration of the AVMP and PW Algorithms.

4. Algorithm Performance

In the current CMIS EDR scheme, the final Atmospheric Vertical Moisture Profile and the Precipitable Water EDRs are obtained by post-processing of the core module outputs, as explained previously. Therefore the performance characteristics of the core module are driving the final EDR performances. Refer to the Core Module ATBD for more details.

4.1. Vertical averaging

Retrieved and "true" EDRs are vertically averaged, in accordance with the SRD requirements, before computing error statistics. The process for vertical layer averaging is illustrated in Figure 4-1. While the process is denoted in terms of the EDR products, the process is the same for truth data.

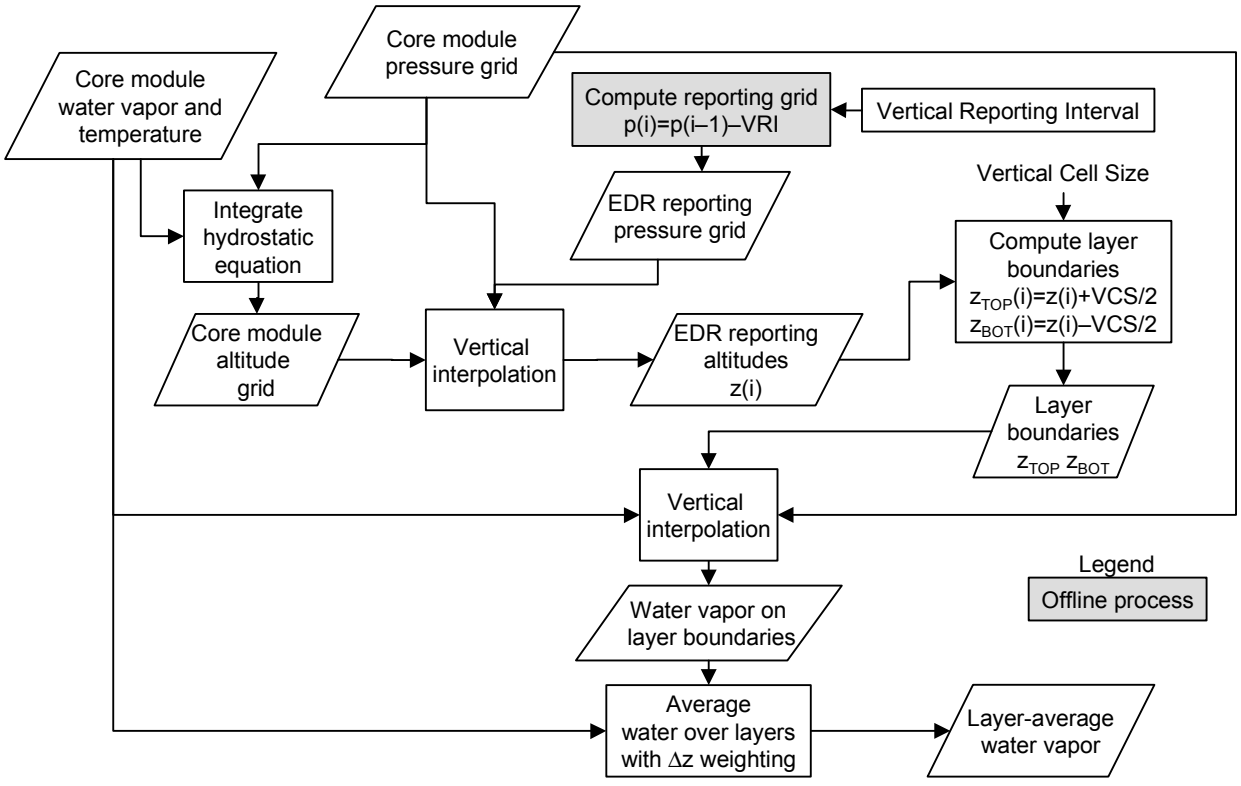


Figure 4-1. Processing flow for vertical averaging for the AVMP EDR validation.

4.2. Test Data Sets

Unless otherwise specified, the test data sets are the ones used to evaluate the performances of the Core Module (refer to the Core Module ATBD for details on the test data sets

characteristics). With respect to "cloudy" performance discussed below, the test conditions used a uniform distribution of liquid water from 0 to 0.5 kg/m^2, which is demanding relative to the range of cloud conditions expected globally.

4.3. Performance Stratification

Above ~500 mb (upper troposphere), the performance in AVMP and PW are similar over ocean and land as the CMIS channels do not see the surface. The clear and cloudy (cloudy corresponds to random CLW with a maximum of 0.25 kg/m^2) performances are presented in Figure 4-2 (ocean) and Figure 4-3 (land) by geographic zones (tropical, midlatitude summer and midlatitude winter, as described in the Core Module ATBD).

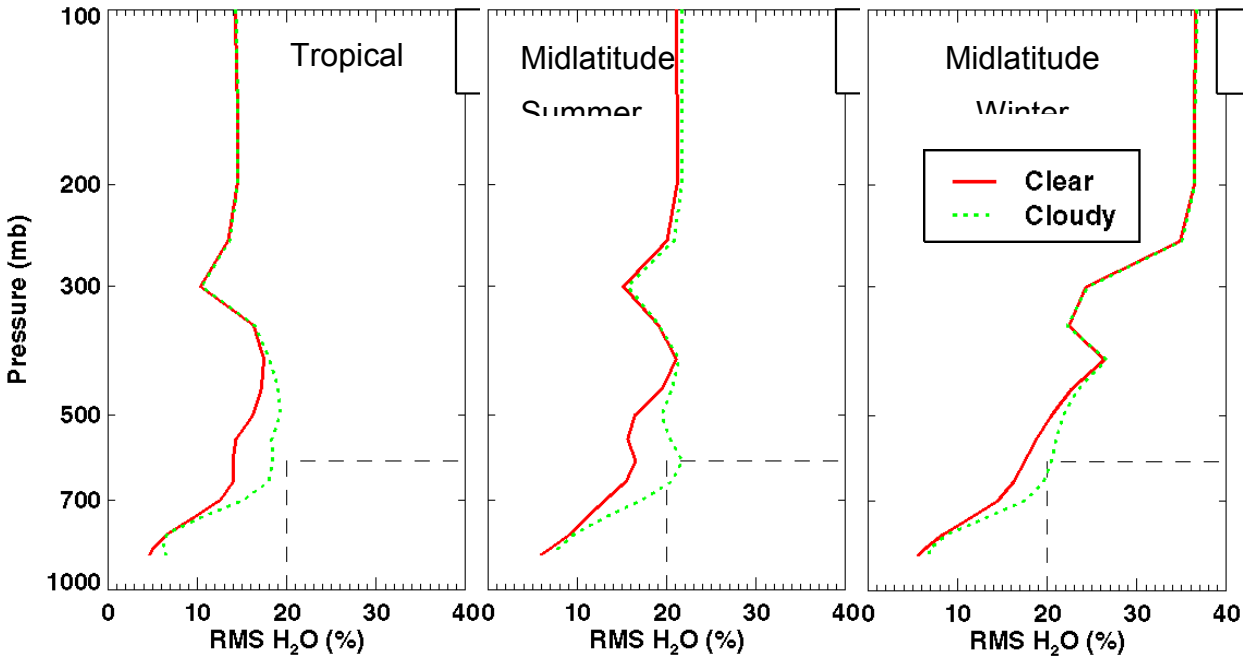


Figure 4-2: Water Vapor Performance Stratification by Geographic Zone – Ocean. Performance was computed with a 15-km CFOV size.

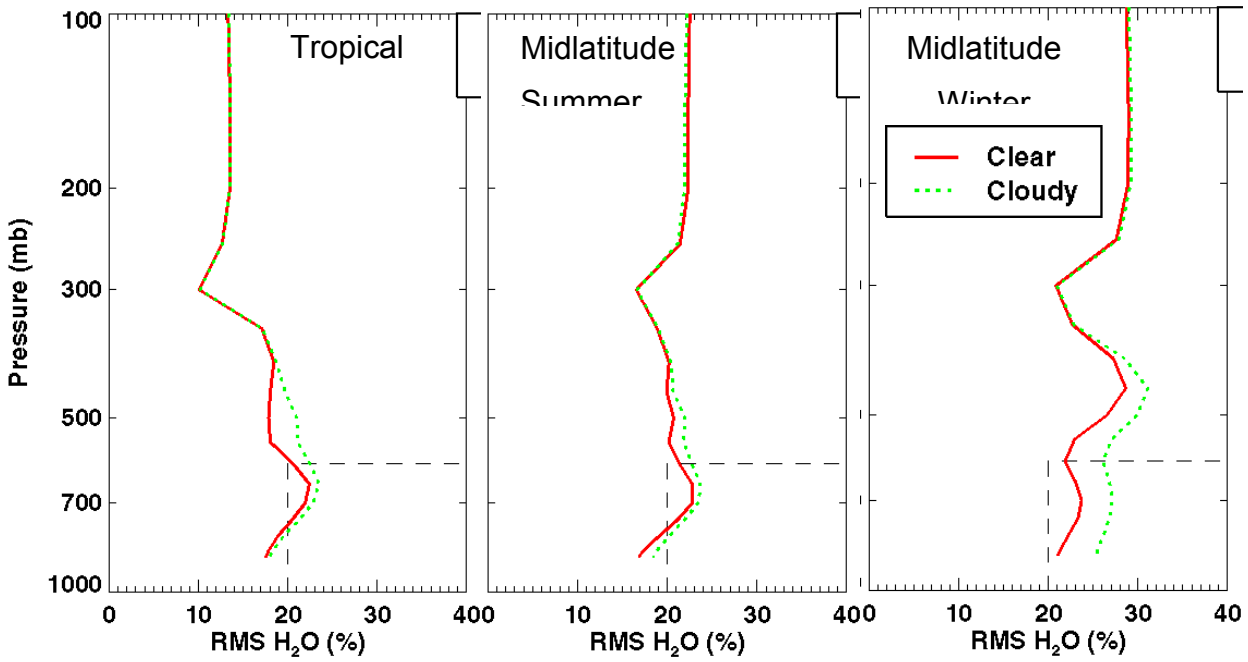


Figure 4-3: Water Vapor Performance Stratification by Geographic Zone – Land. Performance was computed with a 15-km CFOV size.

Water vapor profile performance for cloudy and clear conditions is shown in binned format for a surface type of open shrubland in Figure 4-4. The SRD requirements call for measurement uncertainty performance to be met over the full measurement range. The largest values of 2-km-layer-average mixing ratio in our primary test dataset were about 20 g/kg. Values higher than that are very rare in nature. To test that our algorithm will perform for values up to the top of the measurement range (30 g/kg), we ran retrieval experiments where the "true" mixing ratios were doubled from the values in the database. Retrievals were performed for ocean cases and for a mixture of land sufaces. The performance produced by these experiments (Figure 4-5) did not degrade from what was obtained with the regular datset.

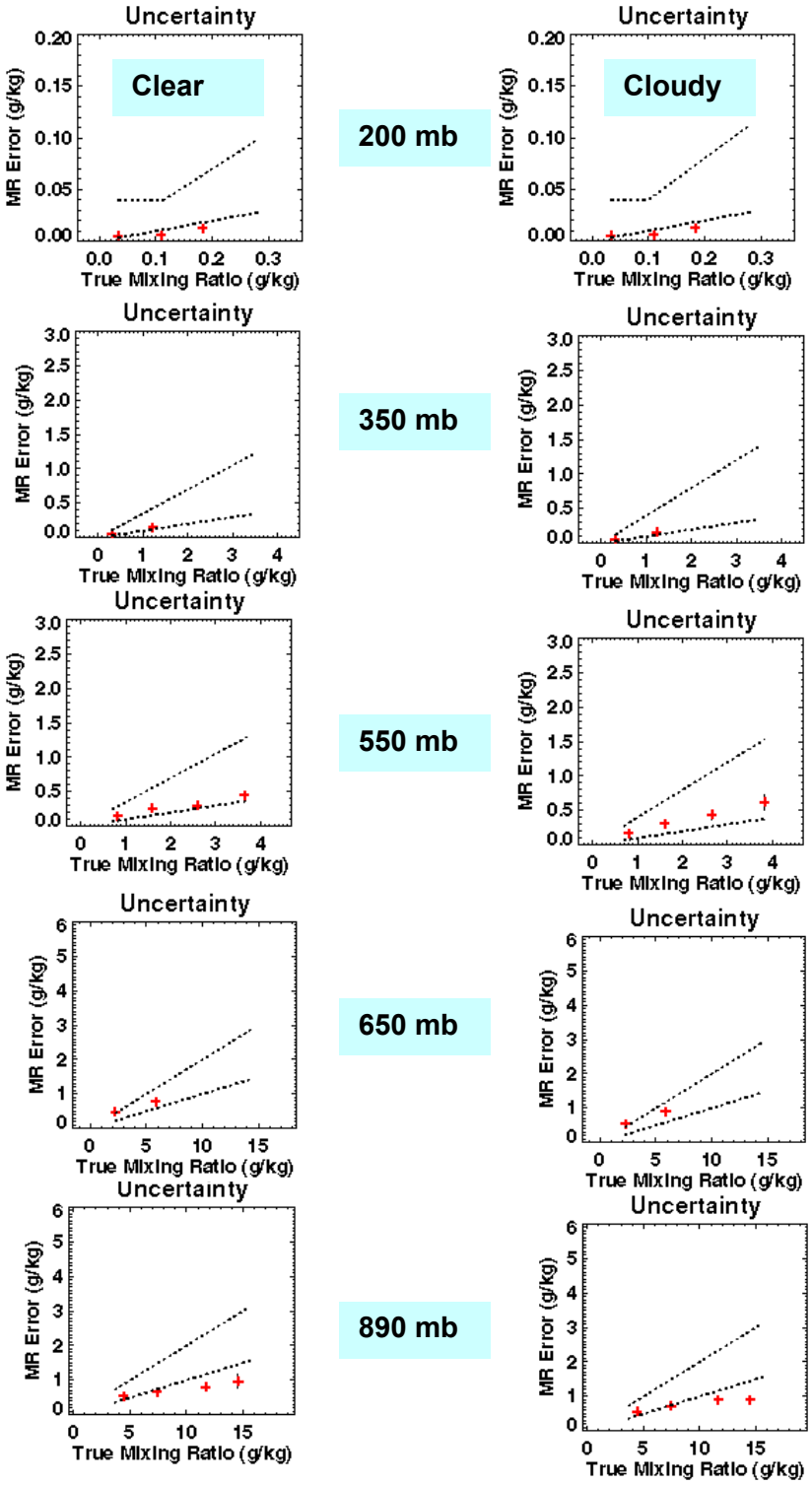


Figure 4-4: Water vapor profile binned performance for the open shrubland surface type, for clear and cloudy conditions. Performance was computed with a 50-km CFOV size. The dotted lines are the threshold and the objective.

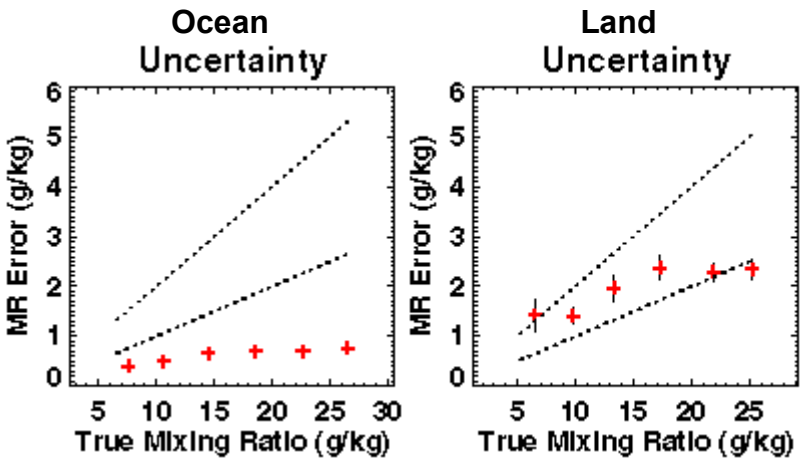


Figure 4-5: Binned AVMP performance for the layer centered on 890 mb for ocean (left) and land (right) test cases where the "true" water vapor was double the value in the standard database. The vertical bars indicate the uncertainty due to sample size, estimated by a bootstrap method. The first land bin had a small sample size.

Performance in the lower troposphere (below ~500 mb) is highly affected by the surface type, as shown in Figure 4-6, Figure 4-7, and . Over land (high surface emissivity) there is a low contrast between the surface and the atmosphere signals, which affects the sensitivity of the channels to water vapor and impairs the ability of the algorithm to separate moisture from clouds (ambiguity between lower troposphere moisture and CLW). Over land, surface emissivity may be highly variable both spatially and temporally, which makes it difficult in many areas of the globe to obtain a good *a priori* knowledge of the surface emissivity.

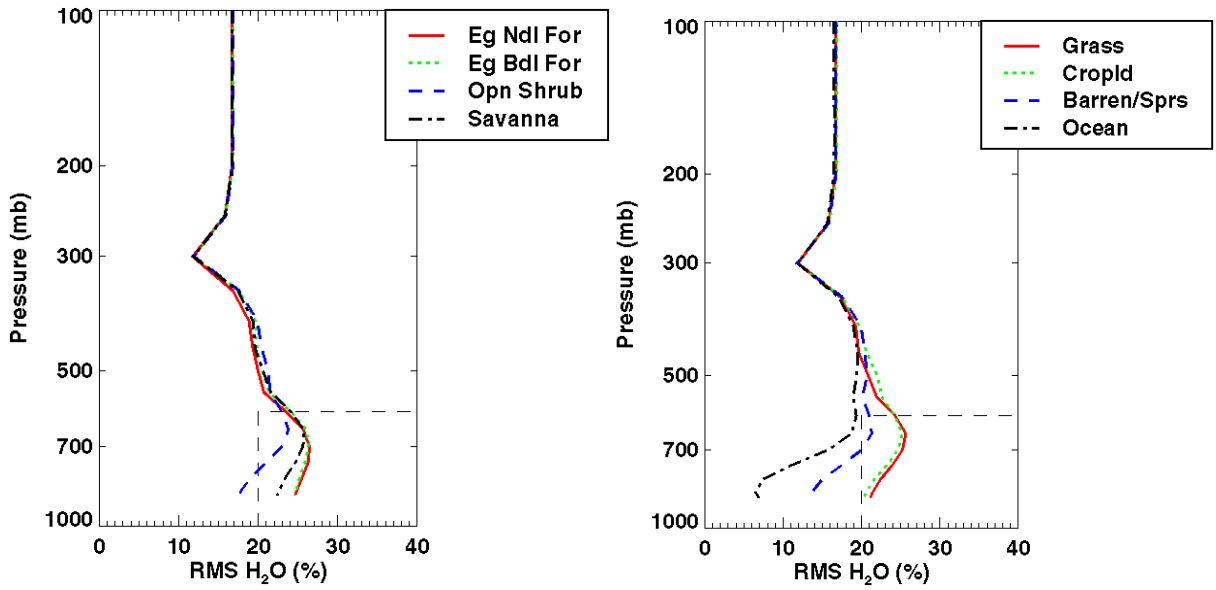


Figure 4-6: Moisture profile performance Stratification by Surface Type – Cloudy. Performance was computed with a 15-km CFOV size.

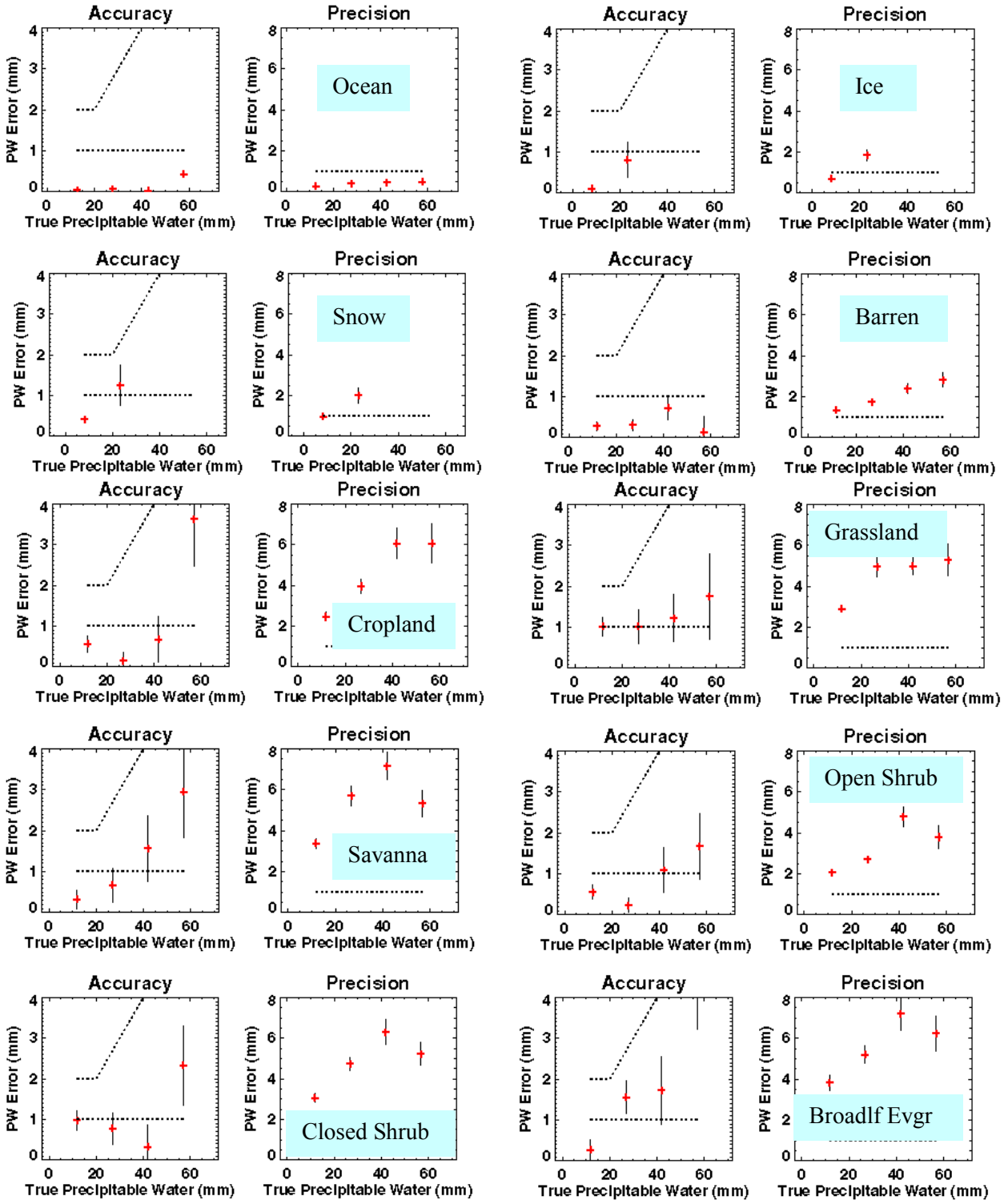


Figure 4-7: Precipitable water binned performance for various surface types, for cloudy conditions. Performance was computed with a 25-km CFOV size.

To first order, water vapor performance over land is driven by the magnitude of 23 GHz H-polarization emissivities. When emissivity is low, water vapor performance meets threshold. As the emissivity of H-polarization approaches 1.0 (such as for vegetated land), water vapor performance degrades rapidly, as shown in Figure 4-8 for PW, due to loss in contrast between atmosphere and surface.

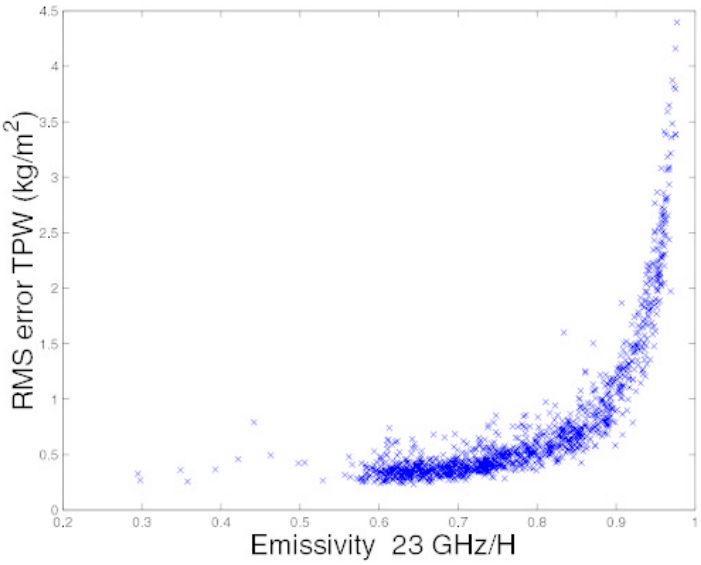


Figure 4-8: Precipitable water errors as a function of surface emissivity.

Table 4-1, Table 4-2 and Table 4-3 summarize the retrieval performance (rms errors) over 15 surface types at 40, 25 and 15 km cell size using global land emissivity constraint. Retrieval test cases used a representative emissivity (closest to the median) from each surface type. 200 retrievals were performed using globally selected atmospheric profiles. PW rms is not stratified according to the PW amount (when the profile is moist, the error should be represented by percent). Stratified rms performance shows that water vapor and cloud error align closely with 18/23 GHz emissivity in horizontal polarization. Figure 4-9 shows the distribution of these land emissivity values over the globe, which are also emphasized in Table 4-4.

Table 4-1: Stratified RMS Performance at 40 km.

Surface Type	18H Emis.	Max. H_2O (%)	PW (kg/m^2)	CLW (kg/m^2)	T_{skin} (K)
Snow and Ice	< 0.80	12.180	1.024	0.047	1.967
Barren or Sparsely Vegetated	0.80 - 0.86	12.519	0.877	0.117	1.664
Open Shrublands	0.86 - 0.90	13.132	1.212	0.154	1.806
Permanent Wetlands	0.90 - 0.94	14.649	1.802	0.173	2.014
Grasslands	0.90 - 0.94	14.450	1.848	0.181	1.816
Croplands	0.90 - 0.94	16.410	2.747	0.183	1.974
Closed Shrublands	0.90 - 0.94	15.750	2.435	0.183	1.890
Savannas	0.90 - 0.94	15.879	2.753	0.177	1.731
Deciduous Broadleaf Forest	0.94 - 0.96	18.283	3.342	0.190	2.018
Cropland/Natural Vegetation Mosaic	0.94 - 0.96	19.290	3.819	0.192	1.988
Evergreen Needleleaf Forest	0.94 - 0.96	19.813	3.956	0.194	2.002
Woody Savannas	0.94 - 0.96	19.599	3.811	0.192	1.894
Mixed Forest	0.94 - 0.96	20.110	4.057	0.198	1.915
Evergreen Broadleaf Forest	0.96 - 0.98	21.263	4.788	0.187	1.950
Deciduous Needleleaf Forest	0.96 - 0.98	23.816	5.798	0.194	1.752

*Tskin threshold at 50 km is 2.5 K

Table 4-2: Stratified RMS Performance at 25 km.

Surface Type	19H Emis.	Max. H_2O (%)	PW (kg/m^2)	CLW (kg/m^2)	T_{skin} (K)
Snow and Ice	< 0.80	12.844	1.360	0.076	2.254
Barren or Sparsely Vegetated	0.80 - 0.86	13.944	1.833	0.121	2.137
Open Shrublands	0.86 - 0.90	14.585	2.264	0.158	2.179
Permanent Wetlands	0.90 - 0.94	15.520	2.960	0.185	2.256
Grasslands	0.90 - 0.94	17.483	3.497	0.183	2.114
Croplands	0.90 - 0.94	12.509	4.420	0.191	2.369
Closed Shrublands	0.90 - 0.94	18.958	4.176	0.192	2.162
Savannas	0.90 - 0.94	19.545	4.196	0.187	2.094
Deciduous Broadleaf Forest	0.94 - 0.96	21.252	4.782	0.199	2.363
Cropland/Natural Vegetation Mosaic	0.94 - 0.96	22.026	5.196	0.198	2.261
Evergreen Needleleaf Forest	0.94 - 0.96	22.161	5.246	0.200	2.291
Woody Savannas	0.94 - 0.96	22.430	5.321	0.198	2.214
Mixed Forest	0.94 - 0.96	23.049	5.507	0.200	2.212
Evergreen Broadleaf Forest	0.96 - 0.98	22.649	5.711	0.199	2.182
Deciduous Needleleaf Forest	0.96 - 0.98	25.587	6.483	0.199	2.033

*TPW threshold at 25 km is 2 kg/m^2 or 10 %

Table 4-3: Stratified RMS Performance at 15 km.

Surface Type	19H Emis.	Max. H_2O (%)	PW (kg/m^2)	CLW (kg/m^2)	T_{skin} (K)
Snow and Ice	< 0.80	14.697	2.157	0.082	2.738
Barren or Sparsely Vegetated	0.80 - 0.86	16.316	3.024	0.119	2.829
Open Shrublands	0.86 - 0.90	17.408	3.535	0.159	2.809
Permanent Wetlands	0.90 - 0.94	20.332	4.428	0.193	2.662
Grasslands	0.90 - 0.94	21.002	5.044	0.188	2.548
Croplands	0.90 - 0.94	23.340	5.422	0.200	2.761
Closed Shrublands	0.90 - 0.94	23.267	5.672	0.199	2.580
Savannas	0.90 - 0.94	24.330	5.912	0.199	2.512
Deciduous Broadleaf Forest	0.94 - 0.96	25.330	6.114	0.209	2.697
Cropland/Natural Vegetation Mosaic	0.94 - 0.96	26.820	6.595	0.211	2.648
Evergreen Needleleaf Forest	0.94 - 0.96	26.204	6.479	0.212	2.568
Woody Savannas	0.94 - 0.96	26.543	6.509	0.211	2.534
Mixed Forest	0.94 - 0.96	27.252	6.819	0.212	2.547
Evergreen Broadleaf Forest	0.96 - 0.98	26.765	6.777	0.213	2.512
Deciduous Needleleaf Forest	0.96 - 0.98	29.046	7.390	0.214	2.396

* H_2O rms threshold is 20 % at 15 km

Table 4-4: Land Emissivity Classes Based on 19 GHz Hpol.

Emissivity	0.98 - 1	0.96 - 0.98	0.94 - 0.96	0.90 - 0.94	0.86 - 0.90	0.80 - 0.86	< 0.8
% Coverage	0.3%	9.6%	21.2%	23.2%	13.6%	16.9%	15.2%

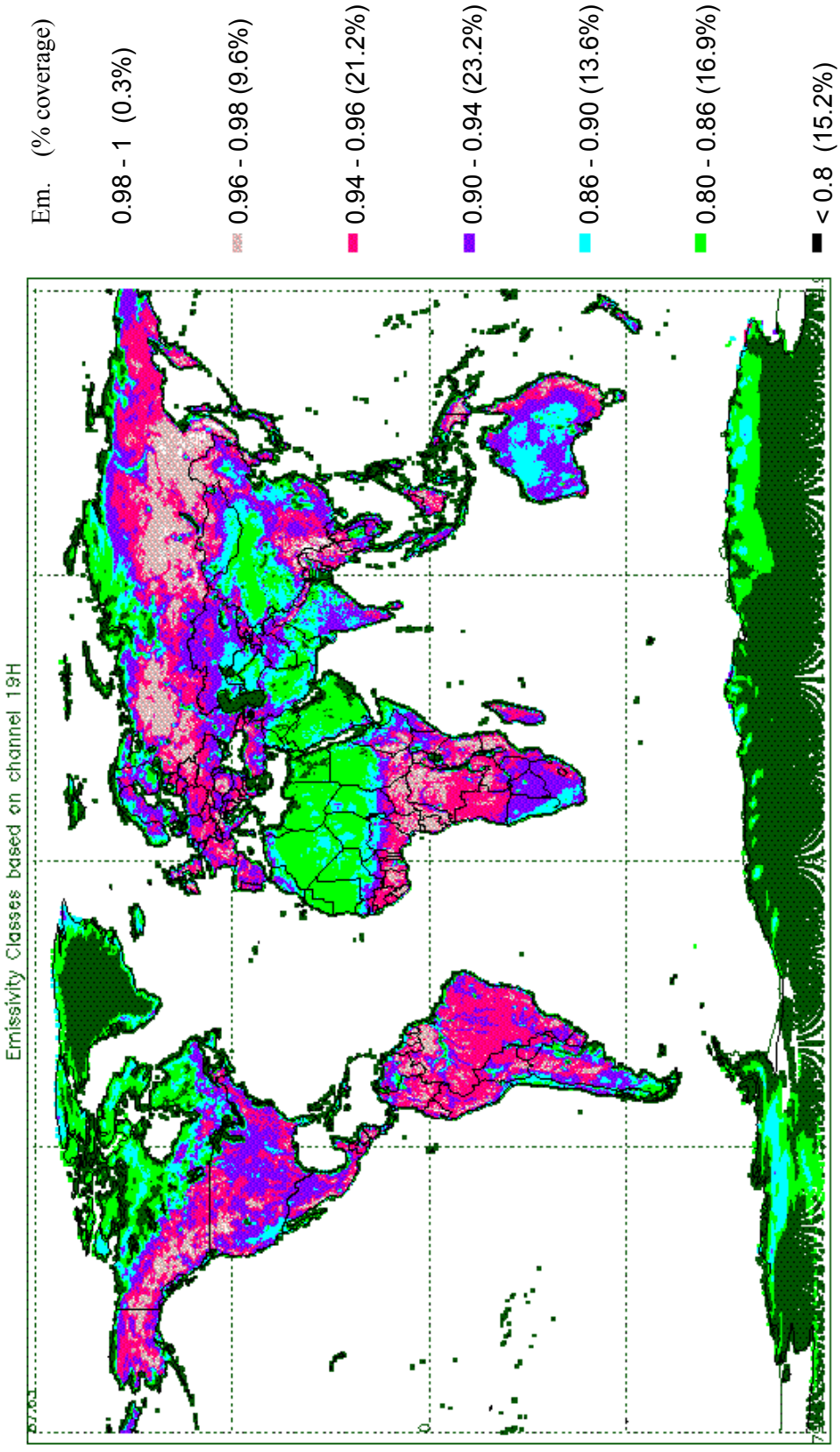


Figure 4-9: Emissivity Classes Based on 19 GHz H.

The information content of PW is low over vegetated land. When surface emissivity is high, radiometric data cannot distinguish between changes in surface emissivity, low-level water vapor and low-level clouds. The radiance signal contribution from cloud liquid water can be compensated, within measurement error, by signal contribution from atmospheric moisture amounts. Figure 4-11 shows the two radiance spectra calculated using the two moisture profiles shown in Figure 4-10. The radiance difference is within noise level for all channels.

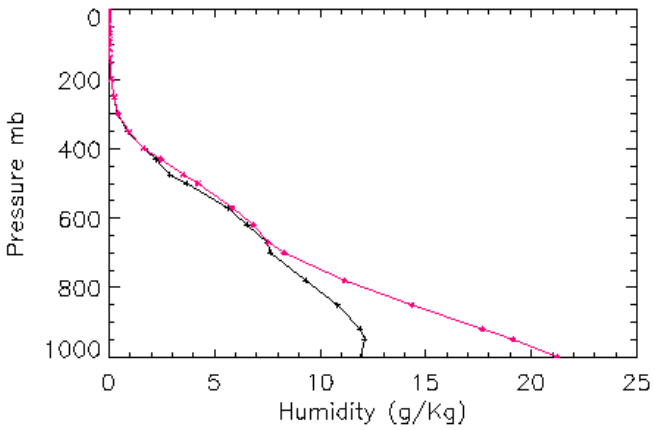


Figure 4-10: Two moisture profiles. The black line is the original profile, the CLW used in simulation is 0.2 kg/m^2. The red line is the retrieved profile (not super-saturated) with CLW=0.0.

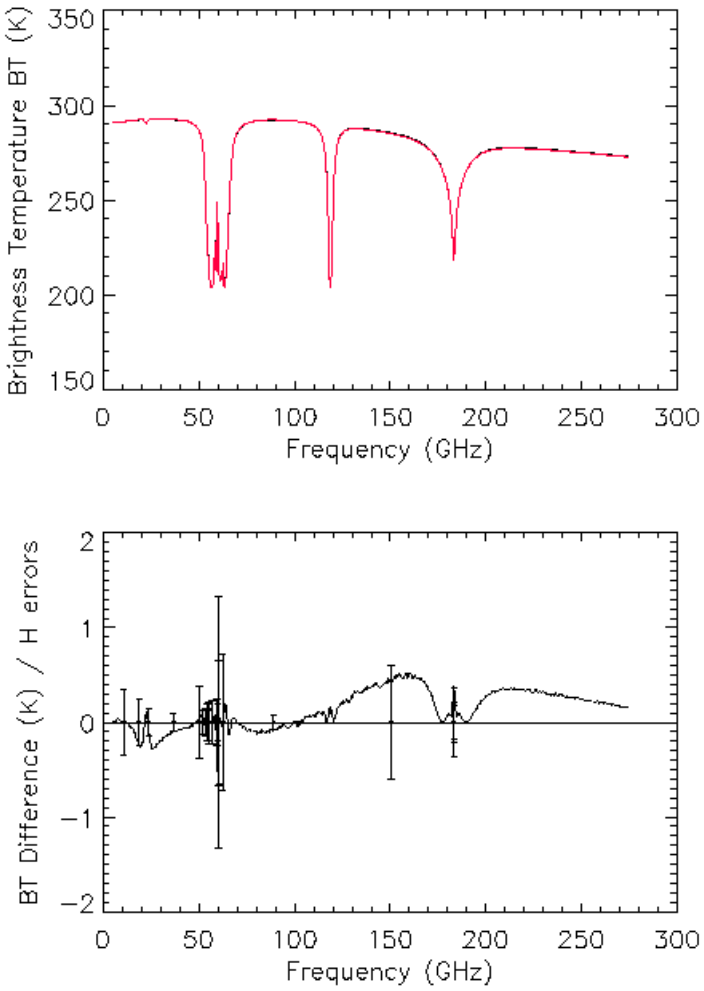


Figure 4-11: Radiance spectra (top) and the difference between them (bottom) corresponding to the moisture profiles shown in Figure 4-10. The vertical bars denote the sensor radiometric noise standard deviation.

4.4. Sensitivity of measurement error to radiometric noise and spatial resolution

The Integration portion of the EDR ATBD (Vol.1) discusses how the core module is executed in a cascade from coarse to fine spatial resolution to minimize computation time and enhance retrieval performance. The performance enhancement comes about because the cascade allows coarse resolution data, which have lower noise by virtue of a larger averaging area, to influence high-resolution retrieval products. The coarse resolution data also include channels at lower frequencies that do not have sufficient spatial resolution to be used directly for high-resolution retrieval processing.

Water vapor profile retrievals are particularly sensitive to changes in radiometric noise that accompany changes in averaging area (CFOV size). Water vapor profiling relies on channels with relatively broad, overlapping weighting functions. Achieving vertical resolution of water vapor features with such data is hampered by data noise. Vertical resolution is, therefore, enhanced by performing retrievals with a larger CFOV size (lower noise; larger horizontal resolution). Improved vertical resolution translates into smaller measurement uncertainties in the 2-km vertical averages over which CMIS data are validated. To the extent that a tradeoff would be necessary between vertical and horizontal resolution, enhanced vertical resolution (and smaller measurement uncertainties) should take precedence. Vertical resolution tends to be the limiting factor when satellite soundings are assimilated for weather analysis and forecasting (Errico, et al., 2000).

The sensitivity of water vapor profile retrieval measurement uncertainty to CFOV size is illustrated in (Figure 4-12). These results were derived from direct execution of the core module at each CFOV size, without using the cascade approach. With the cascade approach applied, the retrievals in areas that are relatively homogeneous horizontally will have vertical resolution and measurement uncertainties about equal to those produced by the 50-km CFOV retrievals.

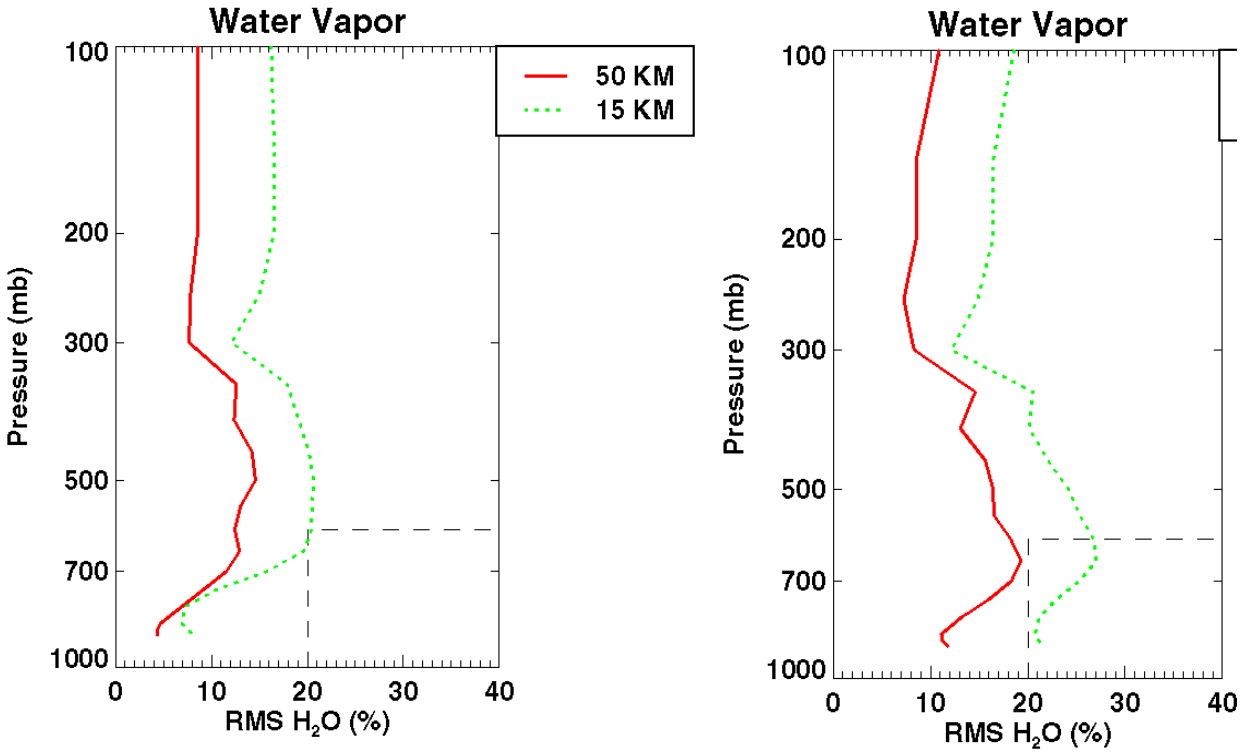


Figure 4-12: The impact on AVMP measurement uncertainty of changing from a noise magnitude characteristic of 50-km CFOV size to one for 15-km CFOV size. The plot at left is for ocean surface and at right is for land surface. Cloudy conditions were assumed in both cases.

4.5. Air mass classification

The core module uses brightness temperatures for sounding channels insensitive to the surface to perform air mass classification prior to retrieval, as discussed in the ATBD for the core module (ATBD Vol. 2). The classification determines which statistical background constraint will be used in a given scene. The impact of classification on water vapor profile retrieval is illustrated in Figure 4-13.

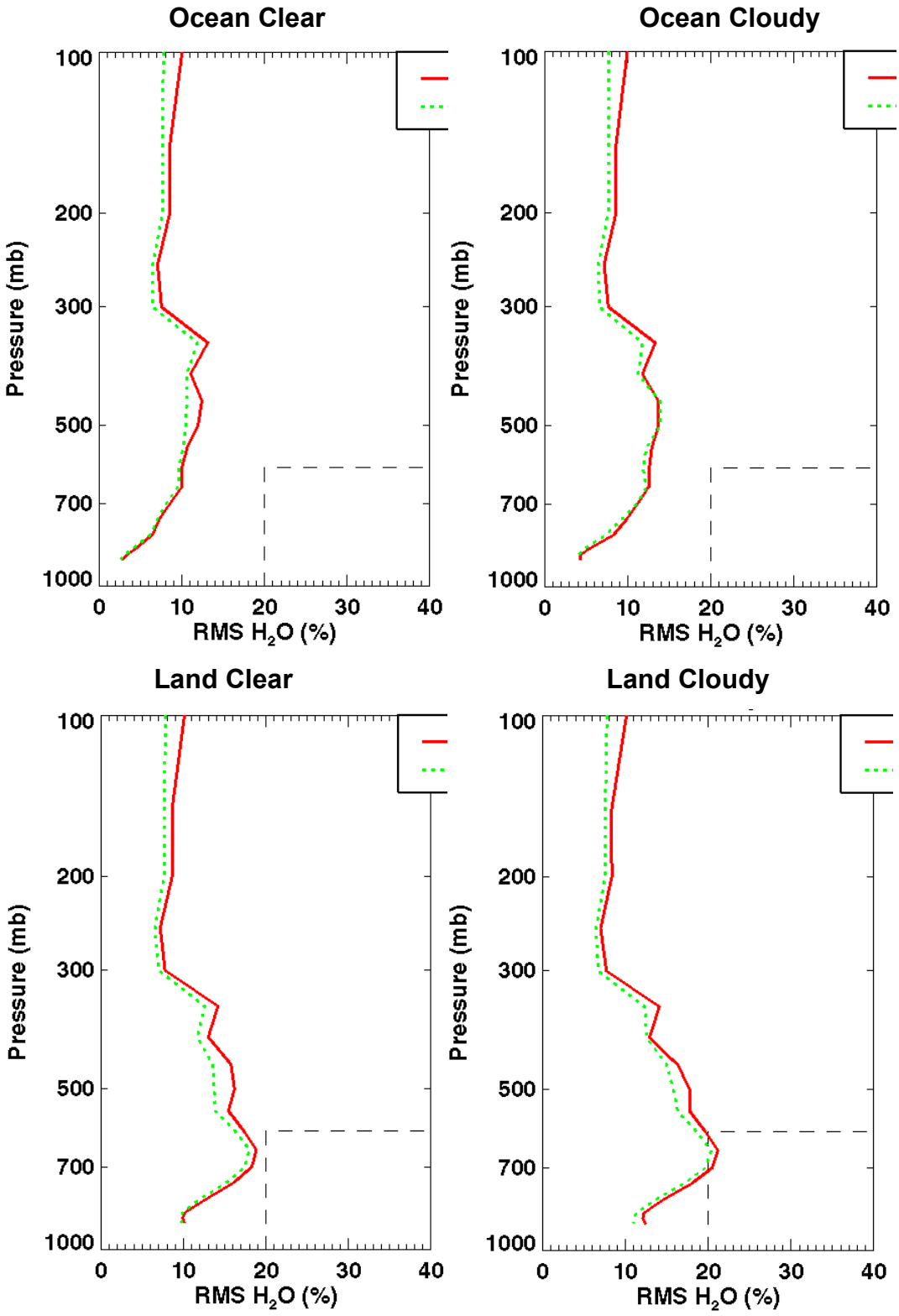


Figure 4-13: Moisture profile performance for global atmosphere background statistics (solid red) and for air-mass classified background statistics (dotted green), for ocean and land surfaces and clear and cloudy conditions. Vertical averaging according to SRD requirements is applied. Performance was computed with a 50-km CFOV size.

4.6. Algorithm treatment of surface emissivity

The core module performs a brightness temperature based pre-classification of surface emissivity, which is focused on identifying the high-emissivity surfaces that are most challenging to water vapor retrieval (ATBD Vol 2., Appendix 6). The major beneficial impact on retrievals for mixed forest surfaces is illustrated in Figure 4-14. About half of these cases were classified as high-emissivity by the algorithm.

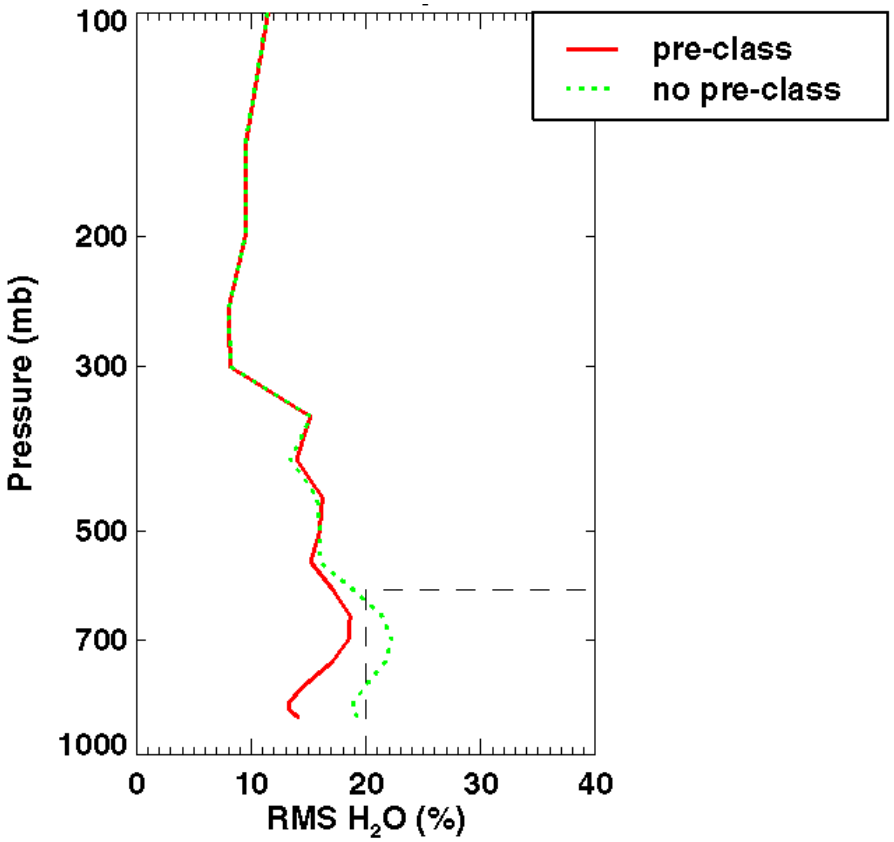


Figure 4-14: Impact of emissivity pre-classification on AVMP retrieval for mixed forest cases. The performance includes a 5% error amplification for cascading from 50-km CFOV to 15 km. A uniform random distribution of cloud liquid water was used ranging from 0 to 0.25 kg/m^2.

The algorithm also has a feature to use a dynamic emissivity database to provide the background estimate. The more stable the emissivity at a location, the better the *a priori* (background) emissivity estimate the dynamic database can provide. That stability is represented in the algorithm in the background error covariance of the emissivity. The influence of the stability on water vapor retrievals for mixed forest surfaces is illustrated in Figure 4-15, which shows a

gradual and substantial improvement in performance going from a global climatological background (0.12 standard deviation) to a stable local background (0.005).

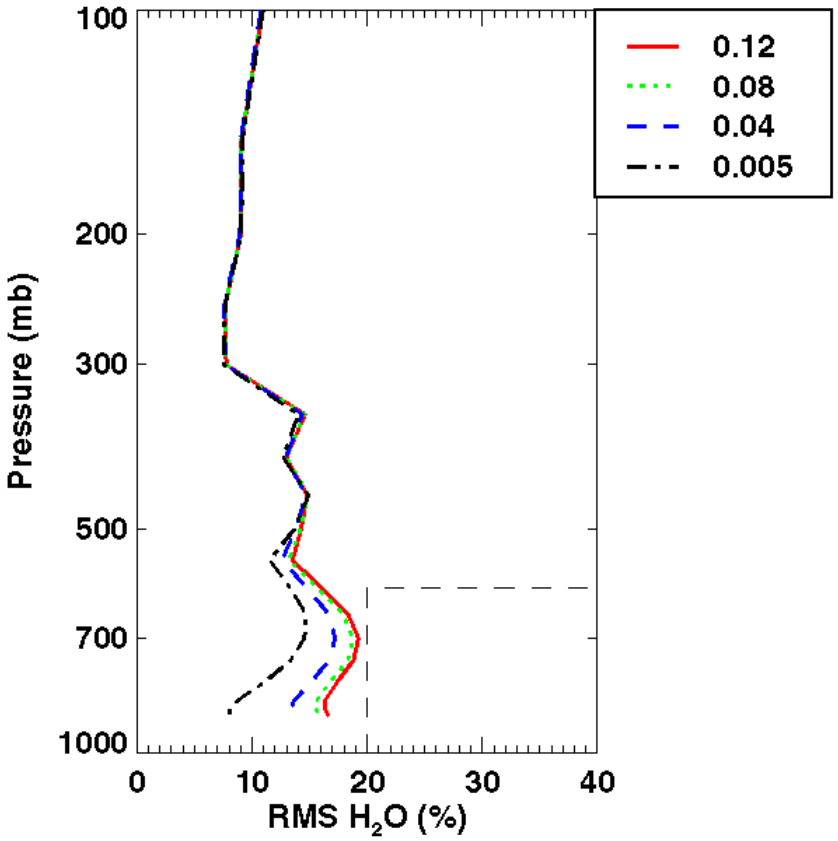


Figure 4-15: The effect of *a priori* uncertainty in surface emissivity on AVMP retrieval performance. The uncertainty is represented as the standard deviation of the 18 GHz H-polarization emssivity. The performance is for clear sky and a 50-km CFOV.

4.7. Sensitivity to retrieval of cloud

The core module typically retrieves cloud parameters simultaneously with the retrieval of water vapor. Cloud water must be retrieved regardless of whether the scene is clear or cloudy unless there is some external information that indicates, in advance of retrieval, that the scene is clear. When external data, such as from VIIRS, are available to identify clear scenes, it is possible to exclude cloud water from the retrieval solution. In such cases, we eliminate a source of ambiguity in the retrieval problem and obtain improved water vapor retrieval, compared with usual case where cloud and vapor must be inferred from the CMIS data. The benefit is particularly significant for high-emissivity surfaces, where the ambiguity is most pronounced. An example is shown (Figure 4-16) for a mixed forest surface type.

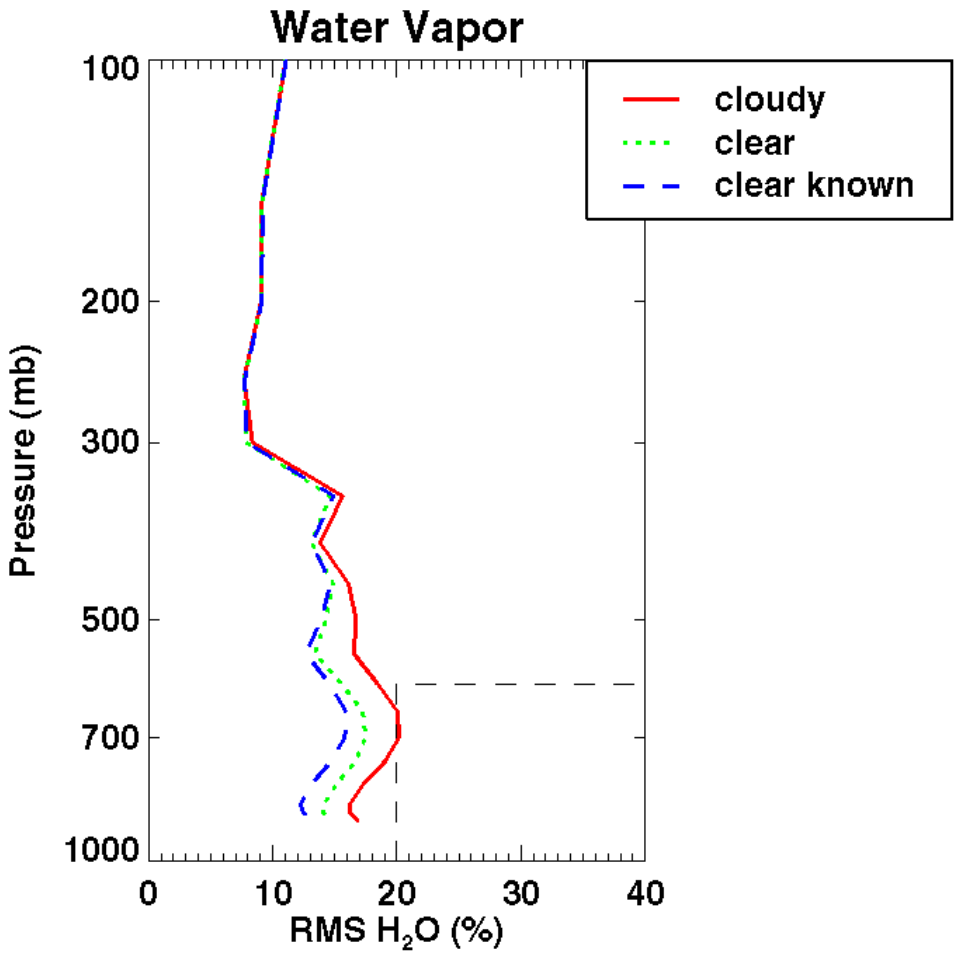


Figure 4-16: Comparison of AVMP for cloudy conditions, clear conditions, and conditions known to be clear from external data. The surface condition is mixed forest, and the CFOV size is 50 km.

Our algorithm approach permits retrieval of water vapor profiles in the presence of most non-precipitating cirrus clouds. This major advance over previous algorithms, which accommodate only thin cirrus, comes from the fact that we can switch to a multiple scattering model. An example of water vapor profile performance (50-km CFOV size) through cirrus clouds is given in Figure 4-17. These results have been obtained with a range of Ice Water Path (IWP) values lower than 1 kg/m^2 and effective mass diameters (D_{me}) up to 800 µm. These conditions extend beyond what would occur without precipitation. The convergence with the non-scattering model (in green dashed) is achieved by relaxing the radiometric error covariance, an idealized situation, so the value of introducing the scattering algorithm is actually underestimated by the difference between the dashed and dotted curves. The results in Figure 4-17 were for a set of test cases

where the cloud microphysics were uniform within the cloud layer. Results for a more demanding set of test cases with a distribution of variations within the cloud layer (Table 4-5) are in Figure 4-18. Results are shown for test cases with cloud tops at 300 mb and at 500 mb. For this set of cases, an algorithm that does not treat scattering cannot retrieve a water vapor profile with errors smaller than the *a priori* (no measurement) estimate. The maximum error amplitude and the pressure level at which it occurs varies with cloud top pressure, and a less sharply peaked error function would occur with an ensemble of cloud top pressures. To provide ensemble performance within CMIS thresholds and to ensure that the 50-GHz channels are not significantly effected, we consider limits on IWP of 0.5 kg/m^2 and on D_{me} of 500 μm to be prudent for reliable water vapor sounding.

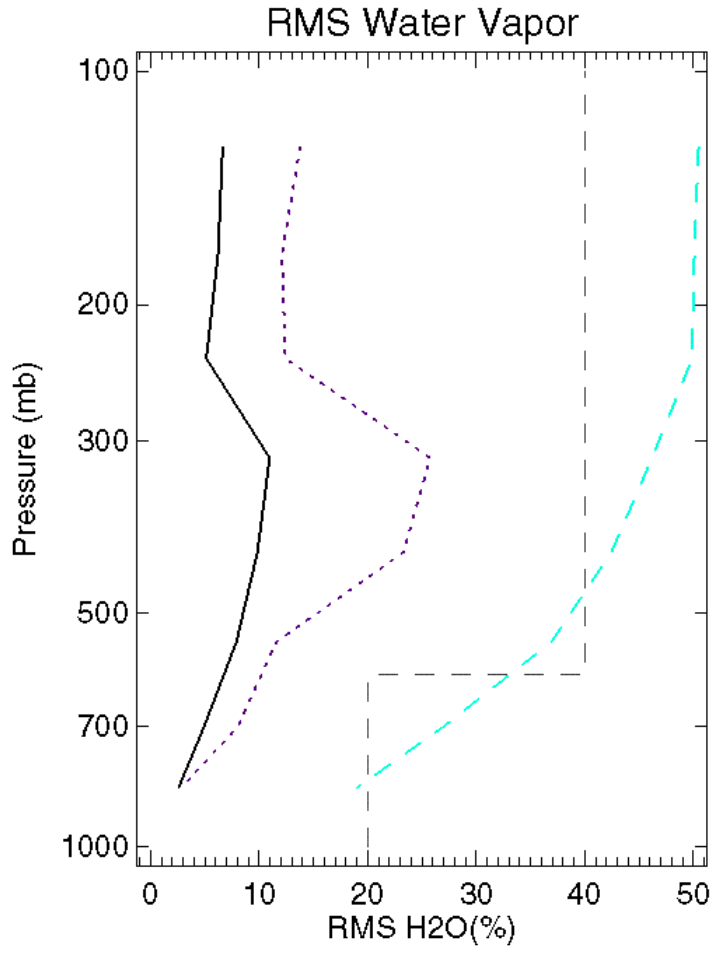


Figure 4-17: Water Vapor Sounding Capability Through Cirrus. Performance was computed with a 50-km CFOV size. The solid curve is performance for clear skies. The dashed curve is for retrieval in the presence of ice cloud with a non-scattering algorithm, while all parameters other than water vapor are held to their true value. The dotted curve is for retrieval in the presence of ice clouds that have tops at 300 mb with the scattering algorithm.

Table 4-5: Variables of ice cloud layer test cases

Cloud layer variable	Range
Cloud thickness	30 to 170 mb
Gamma size distribution width (α)	1 to 7
Decrease in D_{me} from cloud bottom to top	0.5 to 1
Exponent (β) of relationship between D_{me} and IWP (IWP=a D_{me}^{β})	0 to 4

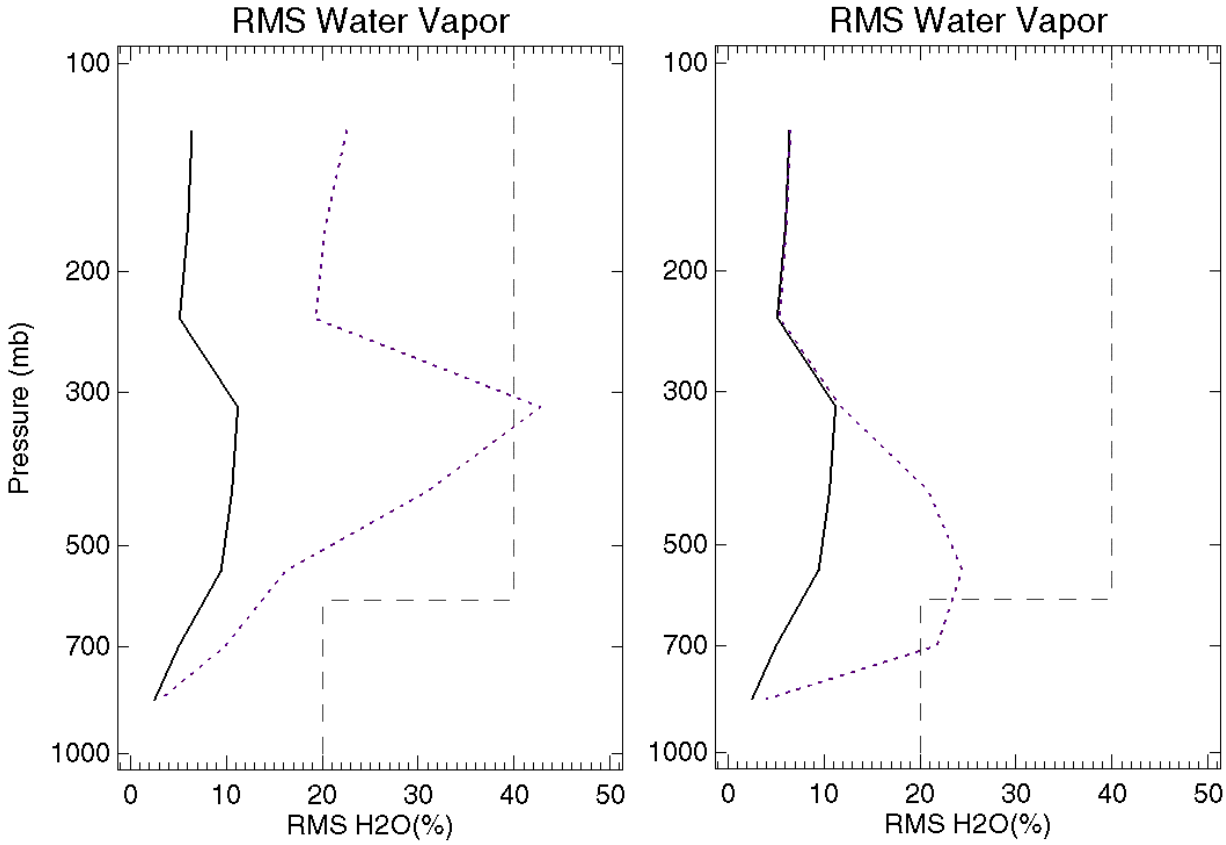


Figure 4-18: Water vapor sounding capability through cirrus with varying microphysics in the layer, as listed in Table 4-5. The left and right plots are for test case cloud tops of 300 mb and 500 mb, respectively. Performance was computed with a 50-km CFOV size. The solid curve is performance for clear skies.

4.8. Additional error sources

Several potential error sources were considered in addition to the error sources covered in the simulations described above. These additional sources are included in the error budget in the following section.

The error budget includes two terms for spectroscopic errors. One refers to the biases between the brightness temperatures CMIS reports and brightness temperatures simulated by applying the core module radiative transfer to data representing the true environmental conditions. It is a residual in the sense that these biases are largely corrected before the water vapor retrievals are performed, using correction factors derived from calibration/validation with ground truth data (Wentz, 1997). The impact of such biases on the AVMP EDR is illustrated in Figure 4-19. Some differences between the CMIS measurements and the model are not sufficiently systematic to be corrected with ground truth data, and the budget includes a separate term for these errors.

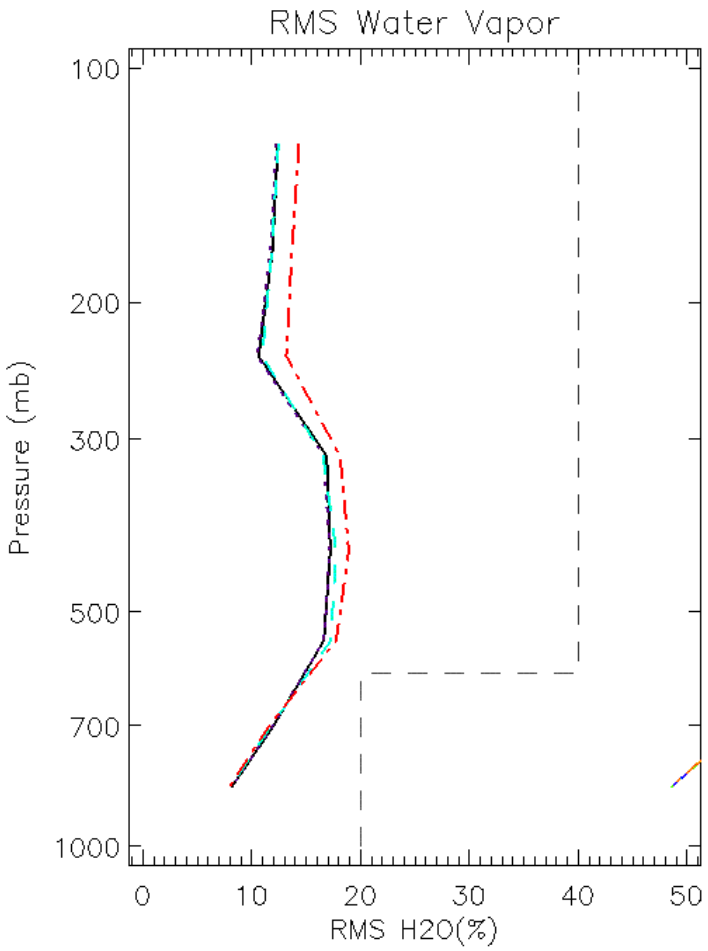


Figure 4-19: The impact of biases in the 183-GHz channels on AVMP retrieval at 15-km CFOV size over an ocean surface. The curves are for no bias (solid), and biases of 0.1 K (dotted), 0.5 K (dashed), and 1 K (dash-dotted).

Sub-field-of-view effects include partial cloud cover within the field of view and differences in cloud between the direct and indirect paths by which radiation reaches the satellite. The direct

path follows the satellite view vector from the surface to the satellite. The indirect path refers to downwelling radiation that is reflected by the surface before being transmitted to the satellite. The effect of partial cloud cover was evaluated by considering the most inhomogeneous case, 50% cloud cover, over the surface where the radiometric data are most sensitive to cloud effects, oceans. We found the impact on water vapor sounding to be negligible, although cloud retrieval skill was affected. To address differences between cloud in direct and indirect paths, we made simulations where cloud was placed in only the direct path and looked at the impact on retrieval performance. An example is shown (Figure 4-20) for a case with a 0.1-kg/m^2 difference between the paths, where the performance is degraded about 3%. These simulations assume a perfectly specular (worst case) ocean surface and a cloud top at 500 mb. We estimate that average cloud variation between the direct and specular paths is 0.05 kg/m^2 and assume that, on a global basis, the effect is reduced from the specular condition by 50%. Another factor arises because liquid clouds are usually lower than 500 mb, and the effect is diminished as the cloud approaches the surface (as the two paths through the cloud become nearly the same). With these adjustments, we estimate a net effect of 0.5%. These evaluations excluded precipitating clouds because precipitating conditions preclude water vapor sounding. The evaluations focused on cloud inhomogeneities, as opposed to other effects, because cloud varies on particularly sort scales. The effect would be present in conditions with cloud only in the indirect path, but those conditions are a small fraction of overall clear (direct path) scenes, so the effect is considered negligible overall for clear conditions.

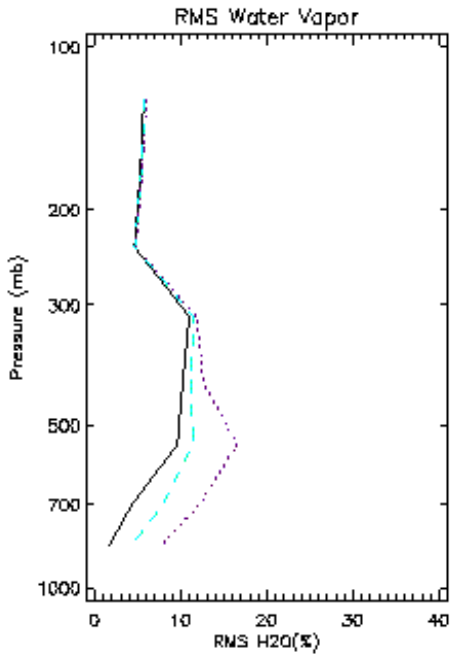


Figure 4-20: Impact of cloud inhomogeneity between the direct path (0.1 kg/m^2) and the indirect path (clear). The cloud top is at 500 mb and the surface condition is ocean. A 15-km CFOV size was used. The dashed curve is for the baseline algorithm and the dotted is for a case where cloud top pressure is not retrieved.

Spatial coregistration errors involve two factors. One is the divergence of two beams that are coregistered at the surface but have different Earth incidence (zenith) angles and, hence, slightly different paths through the atmosphere. An analysis of this factor found no significant effect on water vapor retrieval. The other factor is the uncertainty in the position of each channel's beam in relation to the positions of other channels. This factor was evaluated by considering several types of scene spatial structure that could cause brightness temperatures in a misregistered channel to be different from a correctly registered channel. Effects of cloud, surface emissivity, and surface temperature were considered and the impact on retrieval performance was computed. Details of the analyses are in Appendix TBD of ATBD Vol. 1, Part 1: Integration. For the AVMP EDR, coregistration errors within the requirements flowed to the sensor could cause retrieval error to increase by 3% of the retrieval error obtained without coregistration error. The factor was slightly smaller than 3% over land. For the PW EDR, the error may increase by a factor of 5%.

The channels on CMIS are not all boresighted, so there are time offsets on the order of a few seconds between the views of the various channels. An analysis indicated that the time offsets have no significant impact on retreivals.

The process of vertical registration can introduce error in the vertical moisture profiles through the horizontal interpolations that are involved. We designed the algorithm to minimize those errors. Water vapor retrievals for the 15-km CFOVs are performed at a CFOV spacing of 7.5 km in the along-scan direction so that Nyquist sampling is effectively provided and the interpolation error in that direction is negligible. It is not feasible to produce retrievals complying with Nyquist sampling in the cross-scan direction because the sensor spaces scans 12.5 km apart, although the scans converge near the edge of the swath. The interpolation errors introduced are confined to those associated with features having scales below what can be resolved with 12.5-km spacing and above what is smoothed out in a 15-km average. We estimate a global average of 1% error in the cross-scan direction at center of scan, although the errors could be significantly larger in some isolated conditions. The error is estimated to be 0.6% after averaging over the swath. Analysis of this effect is discussed in Appendix TBD of ATBD Vol. 1, Part3: Gridding.

Errors are introduced by the difference in spatial weighting between the horizontal cells used for validation (uniform averaging over a square) and the composite antenna pattern represented by the CFOV. Analysis of this effect is discussed in Appendix TBD of ATBD Vol. 1, Part2: Footprint Matching and Interpolation.

Surface pressure is required as input to the core module within 2.5 mb uncertainty, driven by the pressure profile EDR requirements. At that level, there is negligible effect of surface pressure error on the water vapor retrievals (Figure 4-21). For uncertainties greater than 5 mb the error becomes significant near the surface, although the errors do not become a significant factor in the maximum error over the surface to 600-mb layer for values up to 10 mb. The effect on precipitable water rms error is less than 0.01 kg/m^2 for uncertainties up to 10 mb.

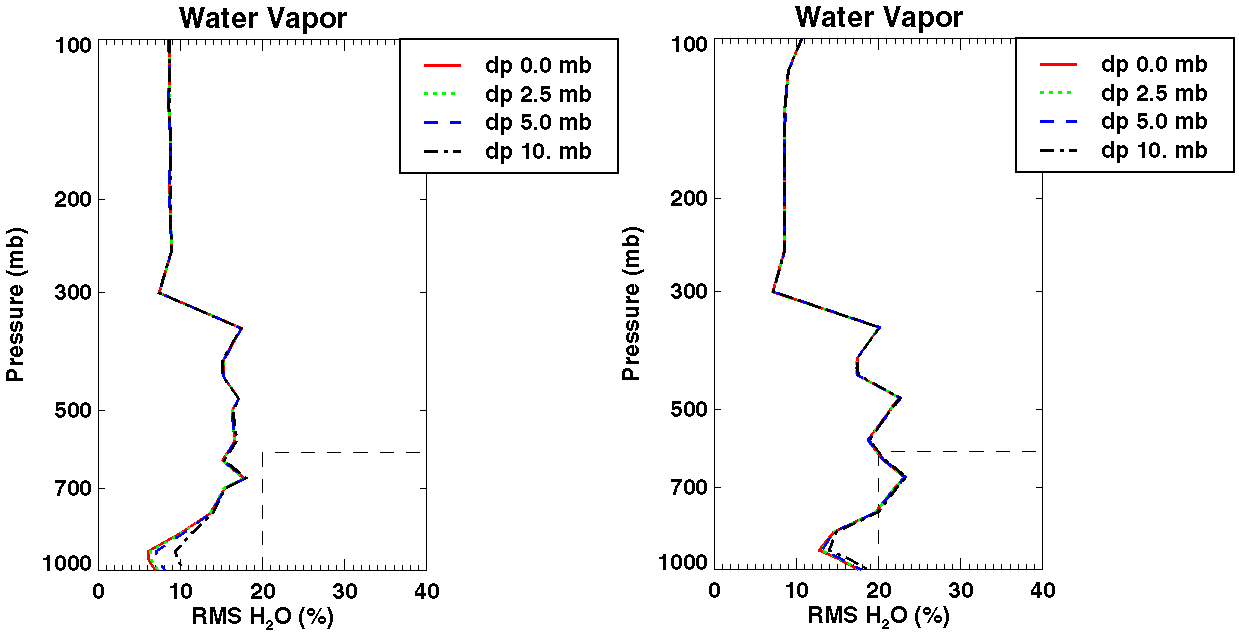


Figure 4-21: Water vapor profile retrieval error as a function of the uncertainty in the surface pressure. The left plot is for ocean surfaces and the right is for land, with clear conditions in both cases. The CFOV size was 50-km and data were not vertically averaged, per the SRD requirements.

4.9. Performance summary

The nominal performance data for the AVMP and PW EDRs are listed in Table 4-6 and Table 4-7, respectively. The key performance metrics are the measurement uncertainty for the AVMP EDR and the measurement accuracy and precision for the PW EDR. The other performance metrics essentially define the measurement domain and the validation conditions under which the key metrics must be met. The error budget discussed below therefore focuses on the key metrics. For AVMP each measurement uncertainty value applies to a range of vertical reporting (pressure) levels. The quoted performance refers to the worst-case pressure level within the range (the performance would appear better if all data within the range were pooled before computing statistics). For both AVMP and PW, performance is computed in bins that span the measurement range and the quoted performance refers to the worst-case bin unless otherwise specified.

Table 4-6: Nominal performance for Moisture Profile EDR.

Para. No.	Description	Threshold	Performance
C40.2.1-1	a. Horizontal Cell Size	15 km	15 km
C40.2.1-2	b. Horizontal Reporting Interval	15 km	15 km
C40.2.1-3	c. Vertical Cell Size	2 km	2 km
	d. Vertical Reporting Interval		
C40.2.1-4	1. Surface to 850 mb	20 mb	20 mb
C40.2.1-5	2. 850 mb to 100 mb	50 mb	15 mb
C40.2.1-6	e. Horizontal Coverage	Global	Global
C40.2.1-7	f. Vertical Coverage	Surface to 100 mb	Surface to 100 mb
C40.2.1-8	g. Measurement Range	0 - 30 g/kg	0 - 30 g/kg
	h. Measurement Uncertainty (expressed as a percent of average mixing ratio in 2 km layers)		
	Clear		
C40.2.1-9	1. Surface to 600 mb	20 % or 0.2 g/kg	18 % or 0.2 g/kg
C40.2.1-10	2. 600 mb to 300 mb	35 % or 0.1 g/kg	22 % or 0.1 g/kg
C40.2.1-11	3. 300 mb to 100 mb	35 % or 0.04 g/kg	22 % or 0.04 g/kg
	Cloudy		
C40.2.1-12	4. Surface to 600 mb	20 % or 0.2 g/kg	20 % or 0.2 g/kg
C40.2.1-13	5. 600 mb to 300 mb	40 % or 0.1 g/kg	40 % or 0.1 g/kg
C40.2.1-14	6. 300 mb to 100 mb	40 % or 0.04 g/kg	40 % or 0.04 g/kg
C40.2.1-15	i. Mapping Uncertainty	5 km	3 km
C40.2.1-16	j. Swath Width	1700 km	1700 km

Table 4-7: Nominal Performance for Precipitable Water EDR.

Para. No.	Description	Thresholds	Performance
C40.3.3-1	a. Horizontal Cell Size	25 km (TBR)	25 km
C40.3.3-2	b. Horizontal Reporting Interval	25 km (TBR)	12.5 km
C40.3.3-3	c. Horizontal Coverage	Global	Global
C40.3.3-4	d. Measurement Range	0 - 75 mm	0 - 75 mm
C40.3.3-5	e. Measurement Accuracy*	1 mm	Ocean: 1mm Land/ice: 1 mm or 5%
C40.3.3-6	f. Measurement Precision*	2 mm or 10%	Ocean: 1 mm Land/ice: 2 mm or 8%
C40.3.3-7	g. Mapping Uncertainty	3 km	3 km
C40.3.3-8	h. Swath Width	1700 km	1700 km

* According to the interpretations in section 2.2.3

The error budget for the AVMP measurement uncertainty is in Table 4-8. (See also EN #65 response.) The "default core module retrieval error" for cloudy cases incorporates cirrus clouds, as discussed in section 4.7. Many of the analyses on which the error budget is based were performed with direct core module retrievals because, at the time of their execution, the simulation environment was not mature enough to make it economical to perform all retrievals with the cascade. The net effect of the cascade depends on the global distribution of fine-scale spatial structure of water vapor and other environmental variables, which is not known with certainty and is difficult to estimate with currently exiting datasets. The error budget includes the effects of radiometric noise and is an estimate of the performance at 15-km cell size upon execution of the cascade. The performance values were obtained by the approximation that, averaged over the globe, the retrieval products will have errors about 5% greater than for 50-km CFOV direct retrievals. The 5% value was derived by considering the results for highly inhomogeneous scenes discussed in ATBD Vol. 1, Part 1: Integration (about 10-20% difference), and accounting for the fact that most scenes will be considerably less inhomogeneous. If the net effect of the cascade is less beneficial than we have estimated, it may be necessary to produce reports for a horizontal cell size greater than 15-km in order to meet the measurement uncertainty requirements. If the net effect is more beneficial than we have estimated, the measurement uncertainties will be smaller than those we have quoted. For precipitable water, which is reported at 25-km cell size, we estimate a 4% relative error increase from the 50-km performance.

Air mass classification was not implemented in the default retrieval simulations, so it is handled as a separate correction (error reduction of 1%) for the purposes of error budgeting. An additional correction is made for the cloudy results regarding the global distribution of cloud liquid. The "cloudy" simulations used a uniform distribution of cloud liquid water from 0 to 0.5 kg/m^2. When considering nonprecipitating cloudy conditions on a global basis, cell-area average liquid water amounts near 0.5 are relatively rare, so the "cloudy" simulations tend to have larger water vapor errors than would be found on a global average. An adjustment is therefore included to reduce the error by 0.3 times the difference between the "clear" and "cloudy" results.

The row in Table 4-8 denoted "Default core module retrieval error" refers to the error sources incorporated into the default retrieval simulations, including smoothing (null-space) and

radiometric noise. Following that row are several rows (up to "Subtotal") for which the errors are additive. The numbers cited are the added retrieval errors that were found as each error source was individually simulated.

The values in the columns for the "300 to 100 mb" in Table 4-8 are given as the same values as the columns for "600 to 300 mb". In our simulations, the errors were generally smaller in the higher of the layers, but we recognize that the radiosonde data we relied on for simulations may not contain as much variation above 300 mb as occurs in nature, considering the degraded radiosonde performance at low temperatures. We took the conservative approach of assuming the relative (%) errors would not be diminished above the 300-mb level.

The values in the columns for the "Surface to 600 mb" in Table 4-8 were derived by considering the range of surface types summarized in Table 4-9. The overall nominal performance for the surface-to-600-mb layer is derived by taking the net nominal performance for each of the surface types. When computing the nominal performance, we have excluded the cases shaded gray in Table 4-8, where performance within thresholds cannot be met. We have not excluded some categories where the performance is slightly over threshold for the worst-case bin but is generally within threshold and, when combined with other categories, the performance meets threshold. For example, the open shrubland (19 GHz H-pol surface emissivity 0.86 - 0.90) performance is near meeting objectives in most bins and levels, but exceeds threshold for a bin at 650 mb (Figure 4-4). The shaded, excluded cases are addressed in the section on performance under degraded measurement conditions below. The net nominal performance combines the performance for conditions where the dynamic surface emissivity database can be applied ($\sigma(\varepsilon_s)$ ≤ 0.04; a conservative estimation) with those where the algorithm relies on global emissivity constraints with preclassification. The "proportion" column indicates the weighting between the two conditions. The results where VIIRS data are not available to identify cloud-free conditions are shown for reference, but are not included in the nominal performance. An exception is the ocean surface, for which the emissivity database is not applicable and for which VIIRS data are not required. The net nominal performances from all the surface types are finally combined by the frequency of occurrence of each type. The occurrence values in Table 4-4 were adjusted by the factor that oceans compose 71% of the surface of the Earth.

Where an error budget entry is zero, that indicates that the error term is negligible in relation to the other terms, not that the error term is identically zero.

Table 4-8: Nominal error budget for Atmospheric Vertical Moisture Profile Measurement Uncertainty

	Clear			Cloudy		
Term	Surface to 600 mb	600 mb to 300 mb	300 mb to 100 mb	Surface to 600 mb	600 mb to 300 mb	300 mb to 100 mb
Default core module retrieval error	15.6%	19.4%	19.4%	18.7%	36%	36%
Adjustment for cascade from 50 to 15-km HCS	0.8%	1.0%	1.0%	0.9%	1.7%	1.7%
Adjustment for air mass classification	−1%	−1%	−1%	−1%	−1%	−1%
Adjustment for global cloud liquid	0%	0%	0%	−1%	0%	0%
Residual calibration/model bias	0.6%	0.9%	0.9%	0.6%	0.9%	0.9%
Residual unsystematic spectroscopic error	0.6%	0.9%	0.9%	0.6%	0.9%	0.9%
Sub-FOV effects	0%	0%	0%	0.5%	0.5%	0.5%
Channel spatial coregistration error	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%
Channel temporal offset	0%	0%	0%	0%	0%	0%
Surface pressure error	0%	0%	0%	0%	0%	0%
Subtotal	17.2%	21.8%	21.8%	19.9%	39.6%	39.6%
Vertical interpolation	0%	0%	0%	0%	0%	0%
Vertical registration (horizontal interpolation)	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%
Cell mismatch	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%
Net	17.2%	21.8%	21.8%	19.9%	39.6%	39.6%

Table 4-9: AVMP EDR performance for the surface-to-600 mb layer for various surface emissivity conditions (19 GHz H-polarization) and other measurement conditions.

Cloudy	VIIRS	$\sigma(\varepsilon_s) \leq 0.04$	Prop-ortion	Ocean	Land ≤ 0.80	Land 0.80 - 0.86	Land 0.86 - 0.90	Land 0.90 - 0.94	Land 0.94 - 0.96	Land 0.96 - 0.98
N	N	N		14%		21%	21%	22%	23%	26%
N	N	Y				19%	21%	21%	22%	26%
N	Y	N	50%		13%	21%	21%	21%	21%	22%
N	Y	Y	50%		11%	19%	21%	21%	21%	22%
Net Nominal Clear:				14%	13%	20%	21%	21%	21%	22%
Y	N	N	50%	18%	15%	24%	24%	24%	25%	31%
Y	N	Y	50%		13%	23%	23%	23%	25%	28%
Net Nominal Cloudy:				18%	14%	24%	24%	24%	N/A	N/A
% Coverage:				71.0	4.4	4.9	3.9	6.7	6.1	2.8

The differences among the surface types for this metric is not as great as one might expect. Variations with surface type are much greater in the layer below 800 mb than they are in the 600-to-800-mb layer (Figure 4-6). Since the errors tend to be larger above 800 mb, the worst-case values for the full surface-to-600-mb layer are generally drawn from the section, above 800 mb, where the surface type has diminished effect.

For precipitable water, there is a large difference in performance between ocean and land, and we therefore summarize the performance separately for the two primary surface types. Ocean refers to areas free of sea ice. The SRD requirements for precipitable water are not stratified between cloudy and clear cases, and the performance does not depend heavily on cloudiness, so we combine the results. We estimate global occurrence of about 50% clear and 50% cloudy, excluding precipitating areas.

Table 4-10: Nominal error budget for the Precipitable Water EDR

	Ocean		Land/Ice	
Term	Accuracy mm	Precision mm	Accuracy	Precision
Default core module retrieval error	0.15	0.25	2.8% or 0.6 mm	7.6% or 1.6 mm
Adjustment for cascade from 50 to 25-km HCS	0.01	0.01	0.2% or 0.0 mm	0.3% or 0.1 mm
Adjustment for air mass classification	0	–0.01	–0.4% or 0.06 mm	–0.8% or 0.2 mm
Adjustment for global cloud liquid	0	–0.01	0	–0.2% or 0.0 mm
Residual calibration/model bias	0.011	0.008	0.3% or 0.06 mm	0.3% or 0.04 mm
Residual unsystematic spectroscopic error	0.011	0.008	0.3% or 0.06 mm	0.3% or 0.04 mm
Sub-FOV effects	0.001	0.003	0.0% or 0.0 mm	0.1% or 0.0 mm
Channel spatial coregistration error	0.013	0.010	0.4% or 0.08 mm	0.3% or 0.06 mm
Channel temporal offset	0	0	0	0
Surface pressure error	0	0	0	0
Subtotal	0.20	0.27	3.6% or 0.7 mm	7.9% or 2.0 mm
Vertical integration	0	0	0	0
Cell mismatch	0	0.17	0	0.5%
Net	0.20	0.32	3.6% or 0.7 mm	7.9% or 2.0 mm

Table 4-11: Precipitable Water EDR performance for various surface emissivity conditions (19 GHz H-polarization) and other measurement conditions.

Cloudy	VIIRS	$\sigma(\varepsilon_s) \leq 0.04$	Prop-ortion	Ocean	Land ≤ 0.80	Land 0.80 - 0.86	Land 0.86 - 0.90	Land 0.90 - 0.94	Land 0.94 - 0.96	Land 0.96 - 0.98
Accuracy										
N	N	N		0% or 0.1	0% or 0.0	1% or 0.3	1% or 0.4	3% or 0.7	5% or 1	4% or 1
N	N	Y			0% or 0.0	2% or 0.3	1% or 0.3	2% or 0.4	5% or 1	4% or 1
N	Y	N	50%		0% or 0.0	2% or 0.3	0.6% or 0.2	1% or 0.6	5% or 1	4% or 1
N	Y	Y	50%		0% or 0.0	1% or 0.2	0.6% or 0.2	1% or 0.7	4% or 1	6% or 1
Net Nominal Clear:				0% or 0.1	0% or 0.0	2% or 0.3	0.6% or 0.2	1% or 0.7	5% or 1	5% or 1
Y	N	N	50%	0% or 0.2	2% or 0.1	2% or 0.3	2% or 0.5	3% or 1	7% or 1	11% or 1
Y	N	Y	50%		1% or 0.1	2% or 0.2	1% or 0.2	2% or 1	7% or 1	5% or 1
Net Nominal Cloudy:				0% or 0.2	2% or 0.1	2% or 0.3	2% or 0.4	3% or 1	7% or 1	8% or 1
Precision										
N	N	N		1% or 0.2	6% or 0.4	5% or 1.3	5% or 2	7% or 2	10% or 2	14% or 2
N	N	Y			5% or 0.6	4% or 1.0	4% or 2	7% or 2	10% or 2	15% or 2
N	Y	N	50%		5% or 0.4	5% or 1.3	5% or 2	7% or 2	9% or 2	14% or 2
N	Y	Y	50%		4% or 0.3	4% or 1.0	4% or 2	6% or 2	10% or 2	12% or 2
Net Nominal Clear:				1% or 0.2	5% or 0.4	5% or 1.2	5% or 2	7% or 2	10% or 2	13% or 2
Y	N	N	50%	1% or 0.3	7% or 0.5	6% or 1.5	6% or 2	9% or 2	14% or 2	22% or 2
Y	N	Y	50%		4% or 0.4	4% or 1.3	6% or 2	8% or 2	15% or 2	20% or 2
Net Nominal Cloudy:				1% or 0.3	6% or 0.5	5% or 1.4	6% or 2	9% or 2	15% or 2	N/A
% Coverage:				71.0	4.4	4.9	3.9	6.7	6.1	2.8

For the lower troposphere (below 600 mb), the nominal performances for AVMP and PW are maintained over ocean and over land, with only high-emissivity land surfaces in cloudy conditions being subnominal.

4.10. Summary of performance under degraded measurement conditions

Measurement conditions that give rise to subnominal performance are summarized in Table 4-12 and Table 4-13. Conditions where the EDRs cannot be retrieved to within any reliable level of performance are listed in Table 4-14 and Table 4-15. The precipitable water retrieval can be made in very light precipitation over ocean because the lower-frequency channels (<60 GHz) are skillful in that environment and the core module quality control mechanism excludes the higher frequencies when they become significantly affected by precipitation.

Table 4-12: AVMP EDR performance under degraded measurement conditions

Condition	Layer	Measurement uncertainty
Cloudy, and surface 19 GHz H-pol emissivity >0.94	Surface to 600 mb	33%
VIIRS cloud cover not available, and Clear	Surface to 600 mb	20%
Prior-pass CrlS water vapor or VIIRS land surface temperature not available, and Clear	Surface to 600 mb	19%
Prior-pass CrlS water vapor or VIIRS land surface temperature not available, and Cloudy	Surface to 600 mb	21%

Table 4-13: Precipitable Water EDR performance under degraded measurement conditions

Condition	Accuracy	Precision
Cloudy, and surface emissivity (19 GHz H-pol) > 0.96	Same as nominal	25% or 2.5 mm
VIIRS cloud cover not available (ocean not affected)	5% or 1.0 mm	10% or 2.0 mm
Prior-pass CrlS water vapor or VIIRS land surface temperature not available (ocean not affected)	6% or 1 mm	9% or 2 mm

Table 4-14: Excluded measurement conditions for the AVMP EDR.

Condition	Indicator	Layer
Surface pressure uncertainty greater than 15 mb	N/A	Surface to 600 mb
Precipitation	Precipitation observed within 8 km of horizontal cell	All
VIIRS cloud cover or cloud top not available, and moderate ice cloud present	Ice water path ≥ 0.03 kg/m^2 or ice particle D_{me} ≥ 300 μm	All
Heavy ice cloud	Ice water path greater than 0.5 kg/ m^2 or ice D_{me} greater than 500 μm	All

Table 4-15: Excluded measurement conditions for the Precipitable Water EDR

Condition
Surface pressure uncertainty greater than 20 mb
Precipitation rate greater than 1 mm/h within 13 km of horizontal cell over ice-free ocean
Precipitation observed within 13 km of horizontal cell over land or ice
VIIRS cloud cover not available, and ice wather path ≥0.03 kg/m^2 or ice particle Dme ≥300 μm
Ice water path greater than 0.5 kg/m^2 or ice D_{me} greater than 500 μm

5. Algorithm Calibration and Validation Requirements

5.1. Pre-launch

To be completed.

5.2. Post-launch

To be completed.

5.3. Special considerations for Cal/Val

To be completed.

5.3.1. Measurement hardware

To be completed.

5.3.2. Field measurements or sensors

To be completed.

5.3.3. Sources of truth data

To be completed.

6. Practical Considerations

6.1. Numerical Computation Considerations

To be completed.

6.2. Programming/Procedure Considerations

To be completed.

6.3. Computer hardware or software requirements

To be completed.

6.4. Quality Control and Diagnostics

To be completed.

6.5. Exception and Error Handling

To be completed.

6.6. Special database considerations

To be completed.

6.7. Special operator training requirements

To be completed.

6.8. Archival requirements

To be completed.

APPENDIX 1: SOUNDING CHANNELS OPTIMIZATION

This appendix outlines the method and two following appendices discuss some details of implementing and testing the results for the 183-GHz channel set.

1 Optimization Method Overview

- Rodgers (1996) optimization measures based on linear maximum probability formulation operate on the posterior estimate covariance matrix
    - information content (Shannon and Weaver, 1949)
    - degrees of freedom for signal
- Simultaneous optimization of center frequency and bandwidth
    - divide the spectrum into small channel elements (discretize) to compute radiative transfer (Jacobian matrix)
    - combine varying numbers of neighboring elements to create channels with varying center frequency and bandwidth
    - optimizer chooses from pallet of channels with different center frequencies and bandwidths
- Simultaneous optimization over multiple base states
    - base state is retrieval vector used to compute Jacobian
        - geographic zones (latitude, surface type, season)
        - cloud conditions
    - track optimization measures for all states
    - select channels that perform well on all base states
    - prevents developing channel set that performs well in general, but poorly in one environment

2 Optimization Method Procedure

Unfeasible to consider all possible combinations

Use multi-step method to find optimum set

- forward selection initial set of narrow-band channels
- replace channels iteratively until optimization measure ceases to improve
    - options at each iteration:
        - replace one channel with another one that is the same width or wider
        - if two passbands touch each other
            - move the dividing line between the two passbands
            - combine the two into one, and add another passband elsewhere

3 Performance Examples BSS/CMIS vs ATMS 183 GHz Moisture Profiling

ATMS results substituted 5 ATMS 183-GHz channels for 3 BSS/CMIS 183-GHz channels, where each set covered water vapor sounding in the 100–1000-mb range.

The IDR channel set is the baseline. Reliability of results above the 300-mb level is low because of test data limitations

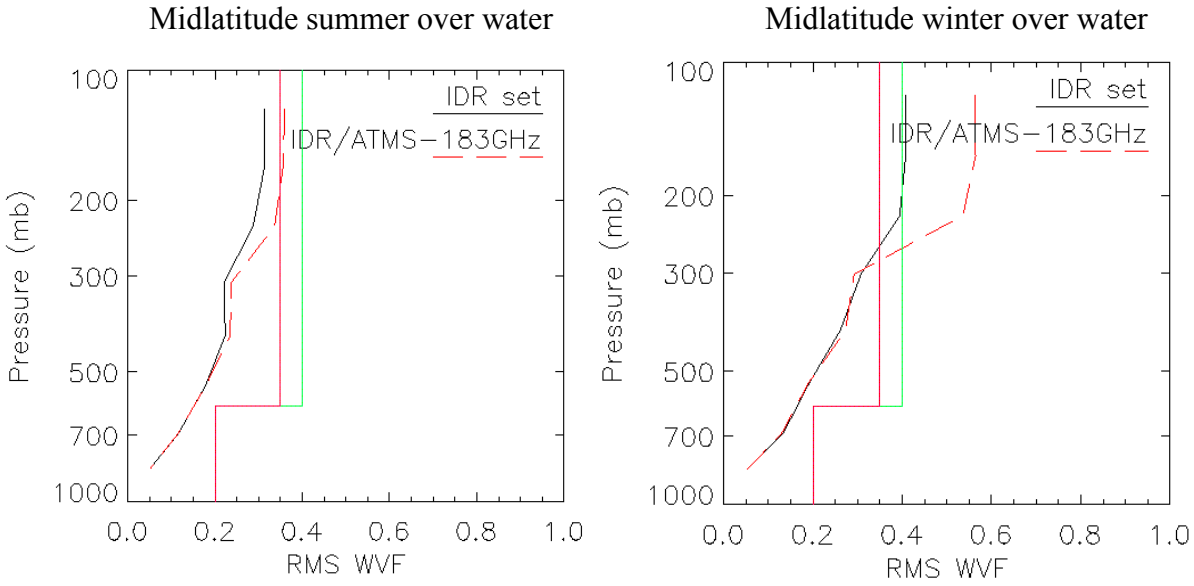


Figure 6-1: Performance Examples BSS/CMIS vs ATMS 183 GHz Moisture Profiling.

APPENDIX 2: DEFINITION OF 183-GHz CHANNEL SET

From CMIS-AIPT-012V2

1 Document Revision

The original version of this document (CMIS-AIPT-012) has been revised to include an appendix regarding NEDT dependence on bandwidth and spectral overlap of channels. The new appendix is cited in the section below on optimization.

2 Optimization

The sounding channels on the 183-GHz water vapor line were defined using the optimization method documented in the briefing package for the 28 June 1998 Technical Interchange Meeting and for the I&A task review of 22 January 1999.

For these channels, optimization was performed simultaneously over nine base states:

1) Tropical ocean
2) Midlatitude summer ocean
3) Midlatitude winter ocean
4) Polar summer ocean
5) Tropical land
6) Midlatitude summer land
7) Midlatitude winter land
8) Polar summer land
9) Polar winter land

No polar winter ocean base state was included because of a lack of sounding data. The base states were averages over stratified soundings from the NOAA88 database.

The palette of spectral intervals from which the optimizer could compose passbands was restricted to double sidebands centered on the 183-GHz water vapor line. The spectrum was discretized in units of 100 MHz. The number of discrete units that could be merged to compose a candidate passband was essentially unlimited, so there was no limit on the width of passbands. Separate optimizations were made where the available spectrum was from ±0.050 to ±8 GHz and from ±0.050 to ±10 GHz. In composing candidate channel sets, the optimizer was precluded

from allowing overlap of passbands from neighboring channels. The justification for the overlap constraint is given in the appendix to this memo. The optimizations took a channel at 150 GHz to be included as a default.

The 183-GHz channels were optimized with respect to water vapor concentration on 21 levels from 94 to 1000 mb. The spacing between levels increased linearly, with respect to the logarithm of pressure, from the bottom upward.

The data noise levels used in the optimizations was from the NEDT budgets supplied by BSS. Single-footprint NEDT values were reduced by averaging factors computed for 15-km cells. Initial optimization experiments were run with the October 1998 NEDT budget and final optimizations were with the February 1999 NEDT budget.

The optimization was specified to assume there were no correlations between errors in brightness temperatures in different channels and no correlations between water vapor concentrations at different levels. The point of the latter assumption was to obtain a channel set that has maximal vertical resolution and does not depend on correlation statistics from any imperfect database.

Optimization scores as function of the number of channels in the 183-GHz set are given in Figure 6-2. The scores are formulated to be indicative of the potential retrieval skill of the channel set. The scores are greater than zero when the number of channel is zero because the horizontal axis labels do not include the 150-GHz channel included as a default. Frame a is from the optimization that produced the SFR-baseline channel set. The scores indicate there is very little benefit to dividing the spectrum among more than 3 channels. Frames b and c are included to show the impact of varying data noise on the trends in optimization scores. Lower noise gives rise to higher optimization scores and a tendency for the scores to level off at a slightly higher number of channels.

The tendency for the optimization scores to level off sharply derives from the relationships between data noise, bandwidth, and radiative transfer. Consider a case where the radiative transfer does not depend on frequency (optical depths are constant with frequency), as is approximately true in atmospheric “windows”. Then changing the passband within the spectrum

would have no effect on the radiative signal. However, the noise for the passband is approximately proportional to the inverse of the square root of the bandwidth. Passbands that do not overlap provide independent measures of the same signal and can be averaged to reduce data noise. When data from two such passbands are averaged, the NEDT of the average value is computed by taking the inverse of the square root of the NEDT for each passband, adding them, and taking the inverse of the sum. If a given portion of spectrum is divided into successively smaller slices, the noise impact of reducing the passband widths is completely offset (within the approximation) by the effect of having more independent samples. If the optimizer was offered a limited portion of the electromagnetic spectrum and had a set number of channels to optimize, it would make the channel bandwidths wide enough to fill the spectrum, and the optimization score would not depend on the number of channels into which the spectrum was sliced.

Along the sides of the 183-GHz line, the radiative transfer does depend on frequency, which is why this spectrum is used for sounding. Passbands are sensitive to different levels of the atmosphere depending on the distance from line center; however, the sensitivity functions (Figure 6-3) are inherently broad with respect to the vertical dimension. Widening the passband furthers the broadening, reducing the independence of separate passbands. However, widening passbands also decreases NEDT. To a degree, the detrimental effect of reducing independence can be offset by the beneficial effect of decreasing NEDT. The optimizer performed this balancing act and found it optimal to widen the passbands enough to essentially fill the available spectrum, whenever the number of channels to work with was three or more (Figure 6-4). So, for more than three channels, the optimization scores could not increase by sampling new spectrum, but only by narrowing the sensitivity functions to enhance independence among channels. Even with infinitesimal bandwidths, the sensitivity functions are quite broad in relation to the depth of the 100–1000-mb layer, so there is not much independence to be gained when going beyond three channels. Furthermore, as bandwidths become infinitesimal, NEDT becomes infinite and the signal-to-noise ratio of each channel goes to zero, so adding channels is clearly not beneficial beyond a certain point.

The issue of independence of passbands is illustrated in Figure 6-5, by considering the sensitivity functions. The 3-channel set of sensitivity functions is shown in Figure 6-5 a, where the middle channel was chosen as the “target” channel. We found the linear combination of the other two

sensitivity functions that minimized the rms difference between the linear combination and the target function. Figure 6-5 b illustrates the same experiment for the 5-channel set, where four channels' sensitivity functions were combined to approximate the fifth. With three channels, there is a big difference between the center channel's sensitivity and the linear approximation, indicating the channels are highly independent. With five channels, there is little difference between the target and the linear combination, indicating the channels have a low level of independence. The *independent* sensitivity of the middle channel is given not by the target curve, but by the difference between the target curve and the linear combination. For five channels, that difference is small.

3 Retrieval Performance Tests

The optimizations were tested in retrieval experiments using the Unified Retrieval algorithm. The test cases were profiles from the NOAA88 database, with no clouds in the "true" profiles. We considered performance for optimized channel sets with 3, 4, or 5 183-GHz channels. The numbers of channels considered were chosen on the basis of the optimization scores and for comparison with the SRR-baseline channel set, which had 5 channels. The experiments were performed with noise corresponding to a 15-km averaging cell and included all channels in the SRR channel set with center frequencies between 18 and 150 GHz, along with the varying sets of 183-GHz channels. Retrieval results were stratified according to geographic environment, according to the same categories listed above.

Sample results are in Figure 6-6, which are typical of the impact of changing channel sets. The uncertainty requirements are plotted, but cannot be directly compared to the curves because the curves were generated without any vertical cell averaging. There is very little impact on performance as more channels are added.

The impact of varying data noise was evaluated by performing test retrievals with reduced noise (Figure 6-7). Retrieval performance improved as noise was reduced by factors of 2 and 4, but channel sets with 3, 4, and 5 183-GHz channels still performed about the same for most geographic groups. The typical impact is illustrated for midlatitude summer land Figure 6-7 a, b, c), while the midlatitude winter land set (Figure 6-7 d, e, f) was exceptional among the groups with respect to the spread among the 3, 4, and 5-channel sets as noise was reduced. These results

imply that the channel count would have little impact on performance even if moderate improvements were made in the sensor noise.

Retrieval tests were also performed with a small data set that, unlike NOAA88, did not rely on extrapolation for water vapor concentrations above 300 mb. Nevertheless, this alternative dataset was not highly reliable for concentrations in the pressure range from 100 to 300 mb. More information about the dataset is in CMIS-AIPT-008. The results for this dataset could not be stratified by latitude or season because of insufficient data, but were stratified by land/ocean (Figure 6-8). The impact of changing the number of channels was small.

We had some concern that the impacts of changing the number of 183-GHz channels might be buffered by smoothness in the vertical profiles of water vapor in the test and training data. Such smoothness implies large correlations between water vapor concentrations at neighboring vertical levels. If the profiles have little vertical structure, it cannot be beneficial to slice the spectrum into more channels, given the way that NEDT depends on bandwidth. We investigated this possibility by perturbing the water vapor profiles in the NOAA-88 data set. The perturbations were done in units of logarithm of mixing ratio, which is the parameter the Unified Retrieval operates on. For each profile and each level, the perturbed value q_l' was

$$q_l' = \frac{(q_l - m_l) + \mathrm{rand}(\sigma_l)}{\sqrt{2}} + m_l \tag{12}$$

where m_l is the mean value for level l, rand(σ_l) is a gaussian random number with mean zero and standard deviation σ_l,, and σ_l is the standard deviation of q_l. With this method of perturbation, q_l' has the same mean and variance as q_l, but the off-diagonal terms of the covariance matrix are reduced by a factor of two.

Retrieval tests were performed with these decorrelated perturbed data, with the adjusted covariance matrix appearing in the retrieval constraints. The perturbed profiles constitute a more difficult retrieval problem than the unperturbed profiles, so the retrieval errors were higher

(Figure 6-9 *vs* Figure 6-6). As with the other tests, there was very little difference in performance depending on the number of 183-GHz channels.

An additional set of retrievals was made to illustrate retrieval performance for a few profiles. For these retrievals, the basic test profile was a mean of the tropical ocean profiles from the NOAA88 dataset. That profile was perturbed by reducing the mixing ratio on the 475-mb level by 25%. The intent was to make a profile feature with fine vertical structure as a challenge to the vertical resolving power of the 183-GHz channel sets. Simulated retrievals were made with the perturbed profile as the "truth". Retrieved profiles are plotted in Figure 6-10 a, with curves for retrievals made using the optimized 3, 4, or 5-channel sets. The retrieved profiles are very similar to each other, and none of the channel sets were able to resolve the perturbation structure. The retrieval errors (relative to the true profile) highlight the differences among the channel sets (Figure 6-10 b). For this case, the errors for the 3-channel set were actually a little smaller at most levels (including the level of the perturbation) than for the 4 or 5 channel sets. Retrieval errors for cases where the perturbation was made at 350 mb and 620 mb are in Figure 6-10 c and Figure 6-10 d, respectively. For the 350-mb perturbation, the 5-channel set did better than the other sets at 350 mb, but the 3 and 4-channel sets did better at some other levels. The results of these retrievals on variations of a single profile are consistent with the results shown above for groups of many profiles, showing no significant performance benefit of increasing the number of channels from 3 to 5.

4 Remaining Issues

There is a critical concern regarding the bandwidths of channels in the optimized sets. These bandwidths are very large in relation to historical designs for sounding at 183 GHz. BSS staff have expressed concern that receiver performance would be degraded for bandwidths as large as the ones in the optimized sets. When computing NEDT values for the optimizations and retrieval tests, the receiver noise figure was taken to be a constant. The concern is that the receiver noise figure may increase for large bandwidths, tending to offset the term in the NEDT equation where bandwidth appears explicitly in the denominator. One option for dealing with this concern is for BSS to demonstrate that receiver performance can be maintained for bandwidths as large as the ones in the optimized sets. Another option is to reoptimize the channels while limiting the bandwidths or including a bandwidth dependence in the receiver

noise figure term in the NEDT equation. The reoptimization would result in channel sets with narrower bandwidths. To maintain retrieval performance, it would be necessary to increase the number of channels so that the entire spectrum around the 183-GHZ line is still being filled with passbands. The two options present a tradeoff between the engineering challenges imposed by large passbands versus the costs associated with carrying an increased number of channels on CMIS.

The method we used for optimizing 183-GHz channels was designed to be robust, and the array of tests performed with the channel sets was designed to avoid excessive dependence on the characteristics of any particular set of test data. Nevertheless, there remains some possibility that the conclusions of the tests may depend on some features of the algorithm or test data. We intend to continue to examine the performance of alternative 183-GHz channel sets as our retrieval algorithm and testbed develop further.

5 Appendix: NEDT dependence on bandwidth and spectral overlap of channels

According to the NEDT budget supplied by BSS, the noise for the passband is approximately proportional to the inverse of the square root of the bandwidth. This dependence derives from the principle of reducing the magnitude of random variations through averaging. If there are N data and each has a random component with a standard deviation of σ, then the standard deviation of a simple average over the data is $\sigma/\sqrt{N}$. A key assumption for this formula is that the variations are uncorrelated. The receiver output contains the geophysical signal with added random perturbations: that is, with added noise. The noise for the entire passband can be viewed as an average over the perturbations among all infinitesimal subbands that comprise the whole passband. The filter for a given channel defines the breadth and weighting of the average. Consider an idealized case where a passband of width B is composed of a simple average of N subbands of width B_0, so $B=B_0N$. If the magnitude of the random perturbation for each subband is σ_0, then the average perturbation over N subbands would have magnitude $\sigma_0/\sqrt{N}$ or, by substituting B/B_0 for N, the average magnitude is $\sigma_0\sqrt{B_0}/\sqrt{B}$. This result is the inverse square-root dependence of NEDT on bandwidth.

If two channels are filtered from a common data stream and the passbands of the two channels overlap spectrally, then the instantaneous perturbations from the geophysical signal are identical

within the region of overlap. Once the data from the overlap region have been averaged into one channel, the potential for noise reduction has been realized. If the same region is averaged into a second channel, then there is no additional net noise reduction for the channel pair, as the noise in the two channels becomes correlated. The only way to achieve noise reduction in regions of overlap is to have independent data streams (feedhorns, receivers, etc.) or orthogonal polarization for the two channels. Requiring orthogonal polarizations at 183-GHz is equivalent to requiring independent data streams because of the lack of practical ortho-mode transducers (OMTs) at these frequencies. Using independent data streams within the 183-GHz channel set is infeasible for CMIS due to weight and power constraints; therefore, passband overlap was precluded in the optimizer.

Acknowledgement: Ideas and information provided by Tom Wilheit, Tom Ellis, Joe Cadwallader, Jamie Kozlevcar, and Andy Stambaugh contributed to the content of the appendix.

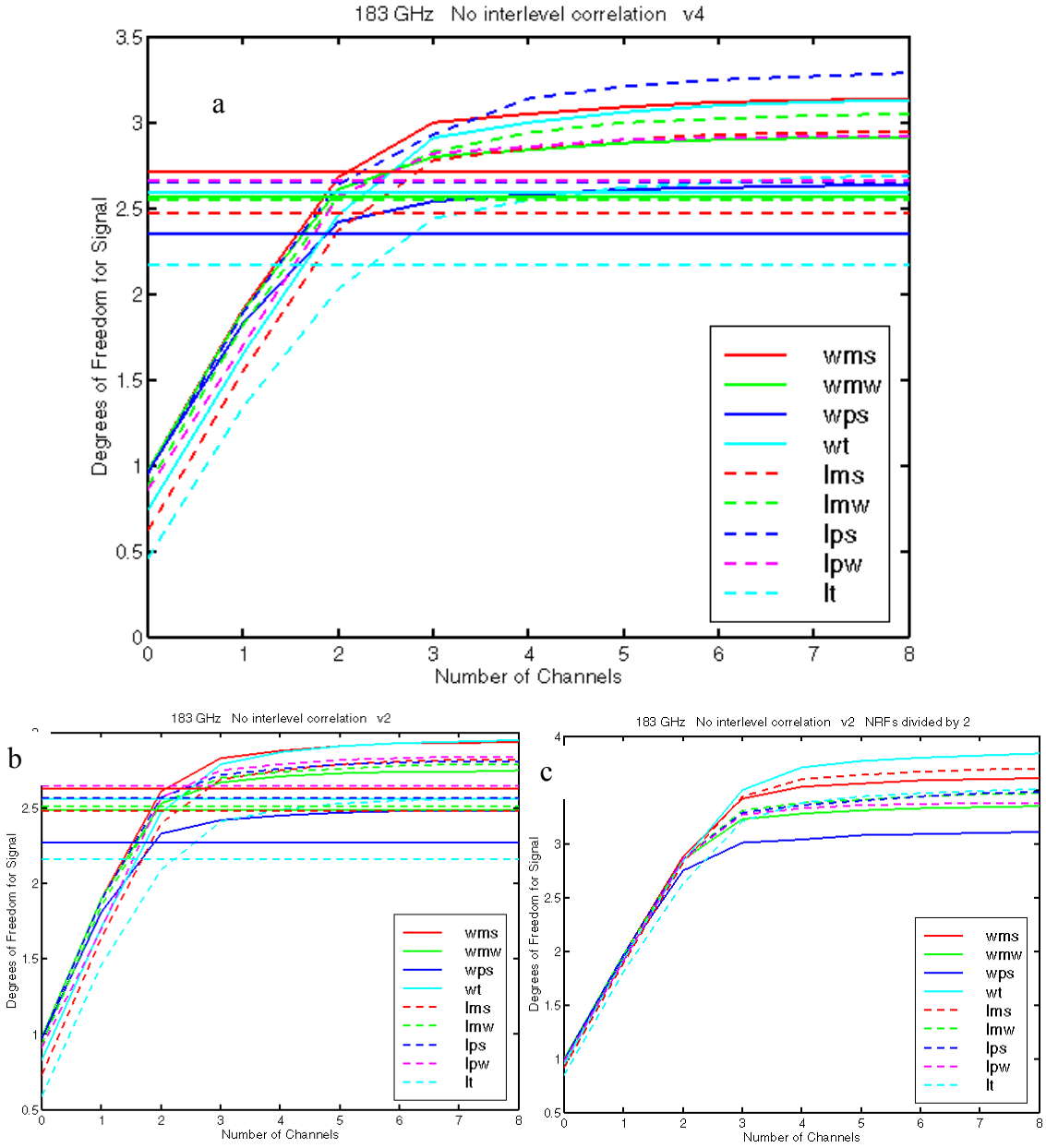


Figure 6-2: Optimization scores as a function of the number of channels in the 183-GHz channel set. For frame a the spectrum in the optimization was 183±10 GHz and the noise was from the Feb. 1999 budget. For frames b and c the spectrum covered 183±8 GHz and the noise was from the Oct. 1998 budget. The radiometric noise was reduced by a factor of two for frame c. The horizontal lines indicate the optimization scores for the SRR-baseline set, which had 5 channels on the 183-GHz line.

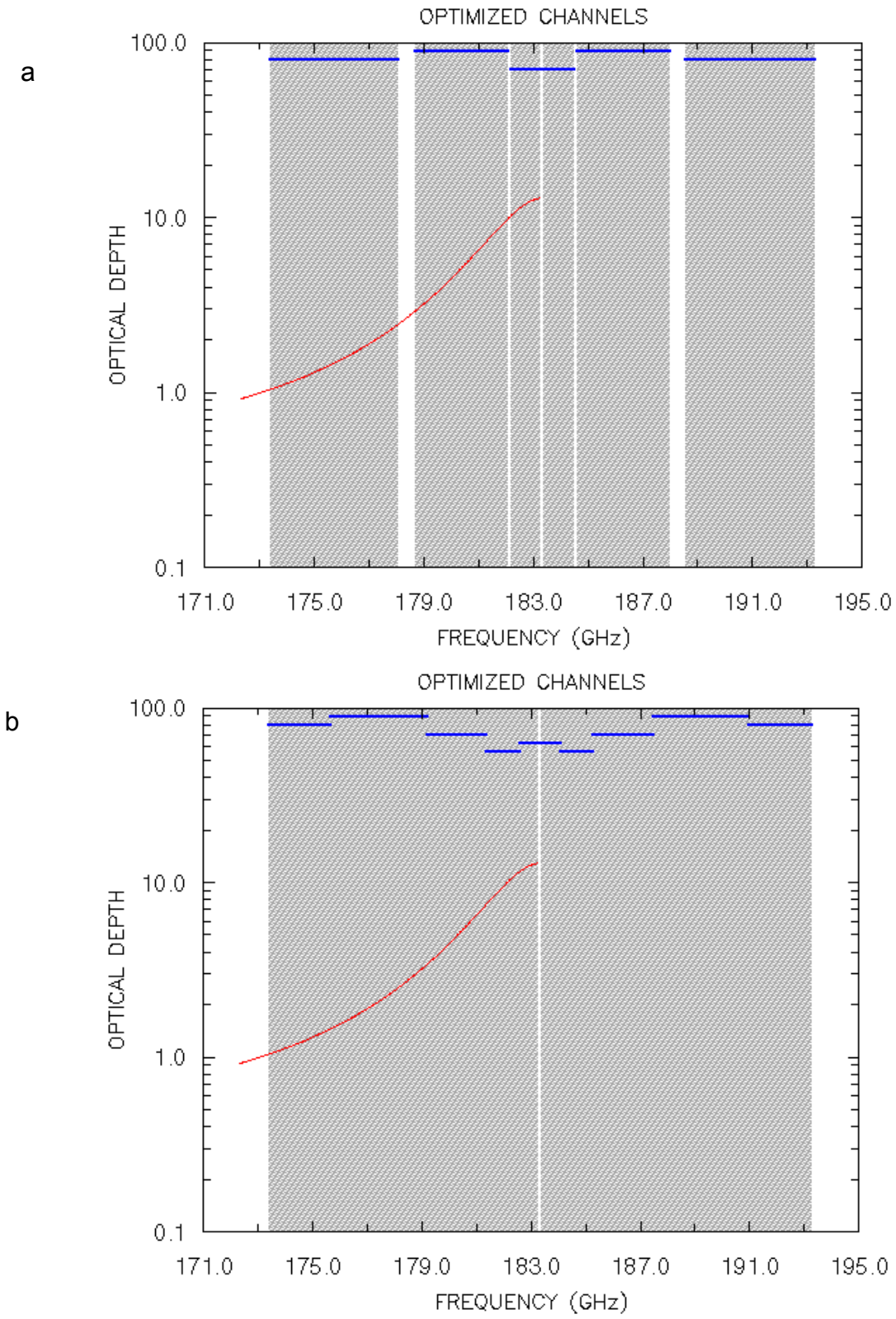


Figure 6-3: Passbands for optimized channel sets with a) three and b) five channels around the 183-GHz line. Passbands are marked by blue lines and shading. Note these are double

sidebands. The red line is the optical depth for double sidebands centered on 183.3 GHz and is symmetrical about the center but is plotted only to the left of the center.

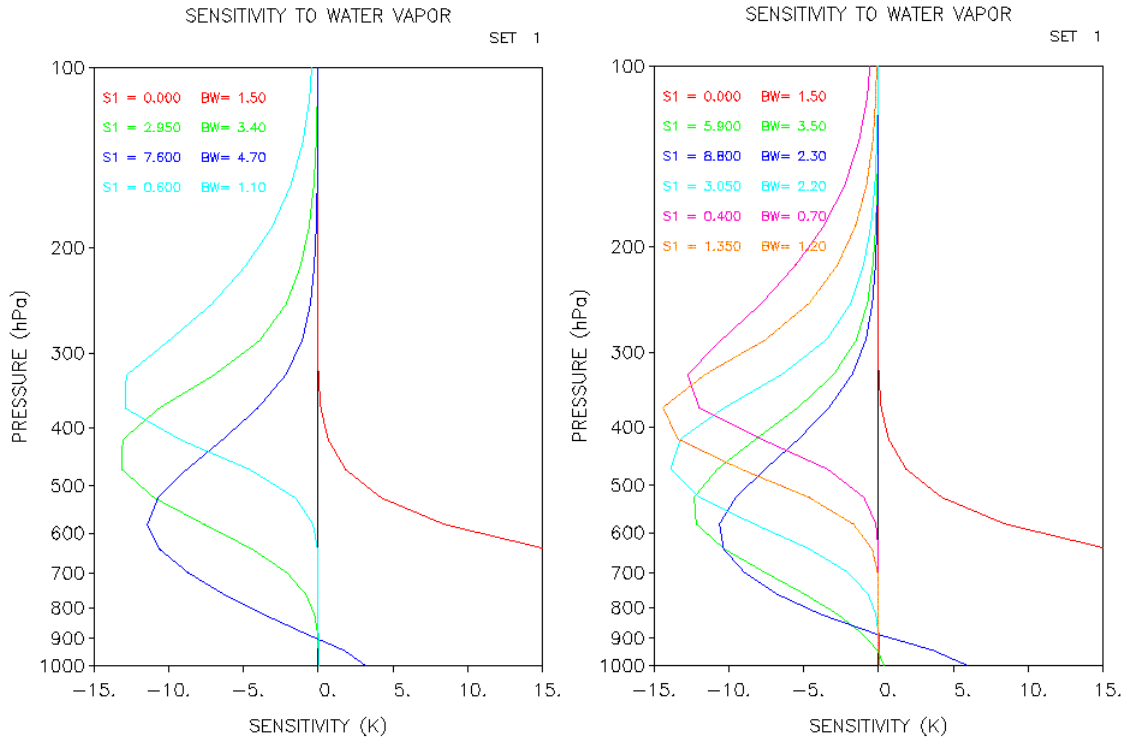


Figure 6-4: Water vapor sensitivity functions for channel sets with a) three and b) five 183-GHz channels. The default channel at 150-GHz is also shown. The functions represent the change in brightness temperature per unit change in water vapor concentration per unit logarithm of pressure. The computations were for the midlatitude summer ocean base state.

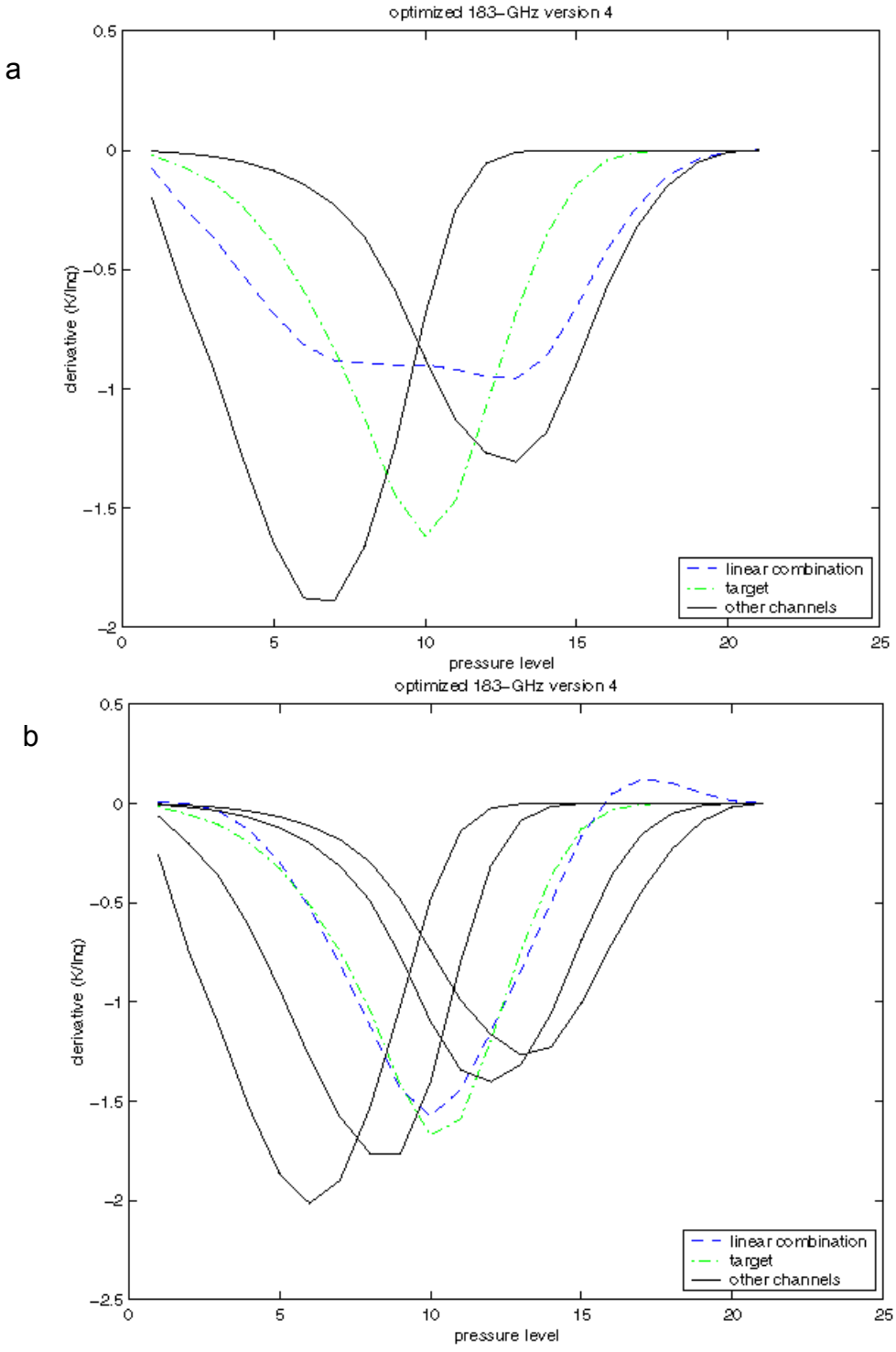


Figure 6-5: Water vapor sensitivity functions (derivative of brightness temperature with respect to logarithm of water vapor mixing ratio) for a) the 3-channel set of optimized 183-GHz channels and b) the 5-channel set. These functions have not been normalized by layer thickness

(in log(p)), unlike those in Figure 6-4, and are plotted as a function of pressure level index rather than pressure. The curves are described in the text.

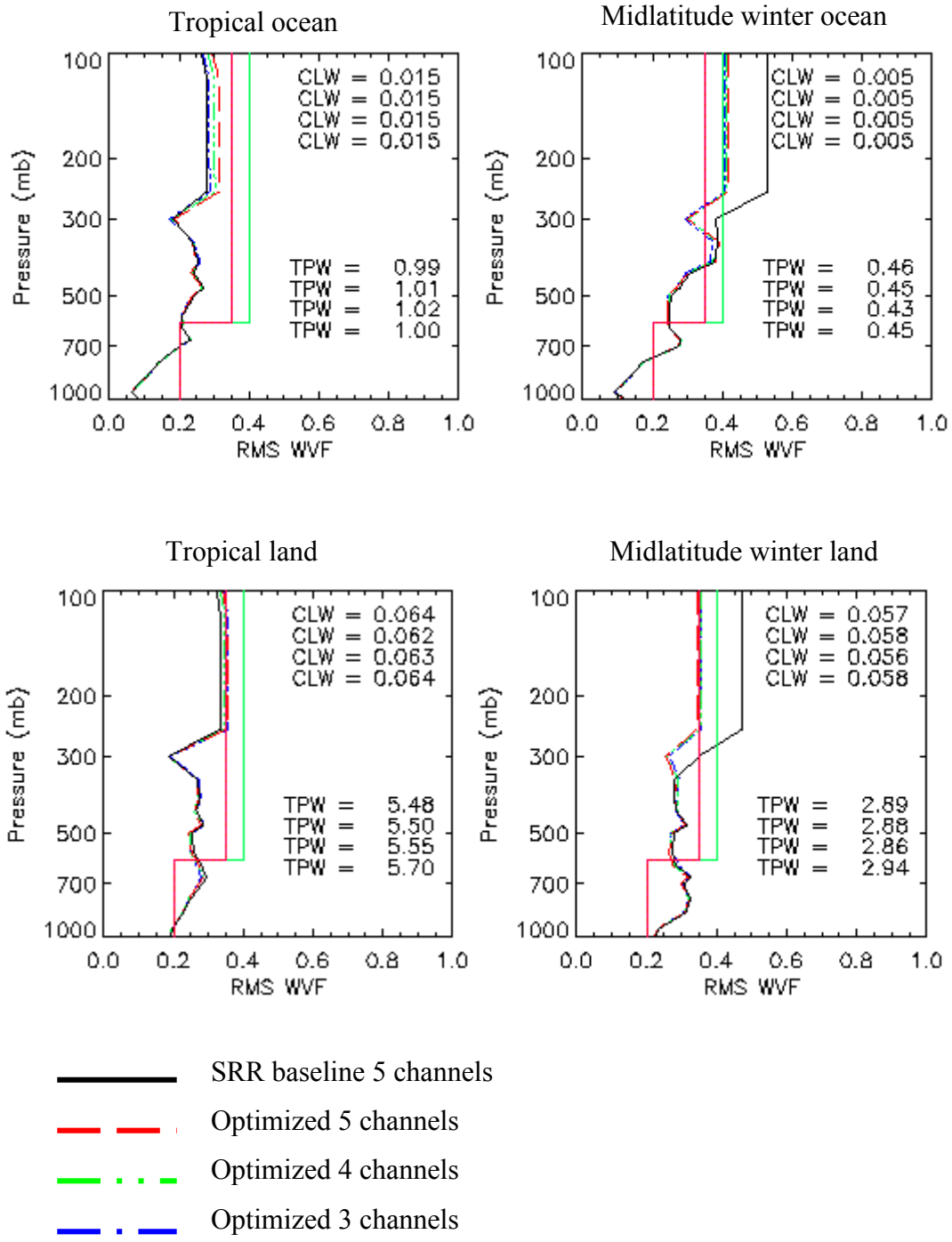


Figure 6-6: Water vapor fractional rms error as a function of pressure. The inset statistics are for cloud liquid water (CLW) and total precipitable water (TPW) at 15-km resolution, given for opt3, opt4, opt5, SRR5 in order from top to bottom.

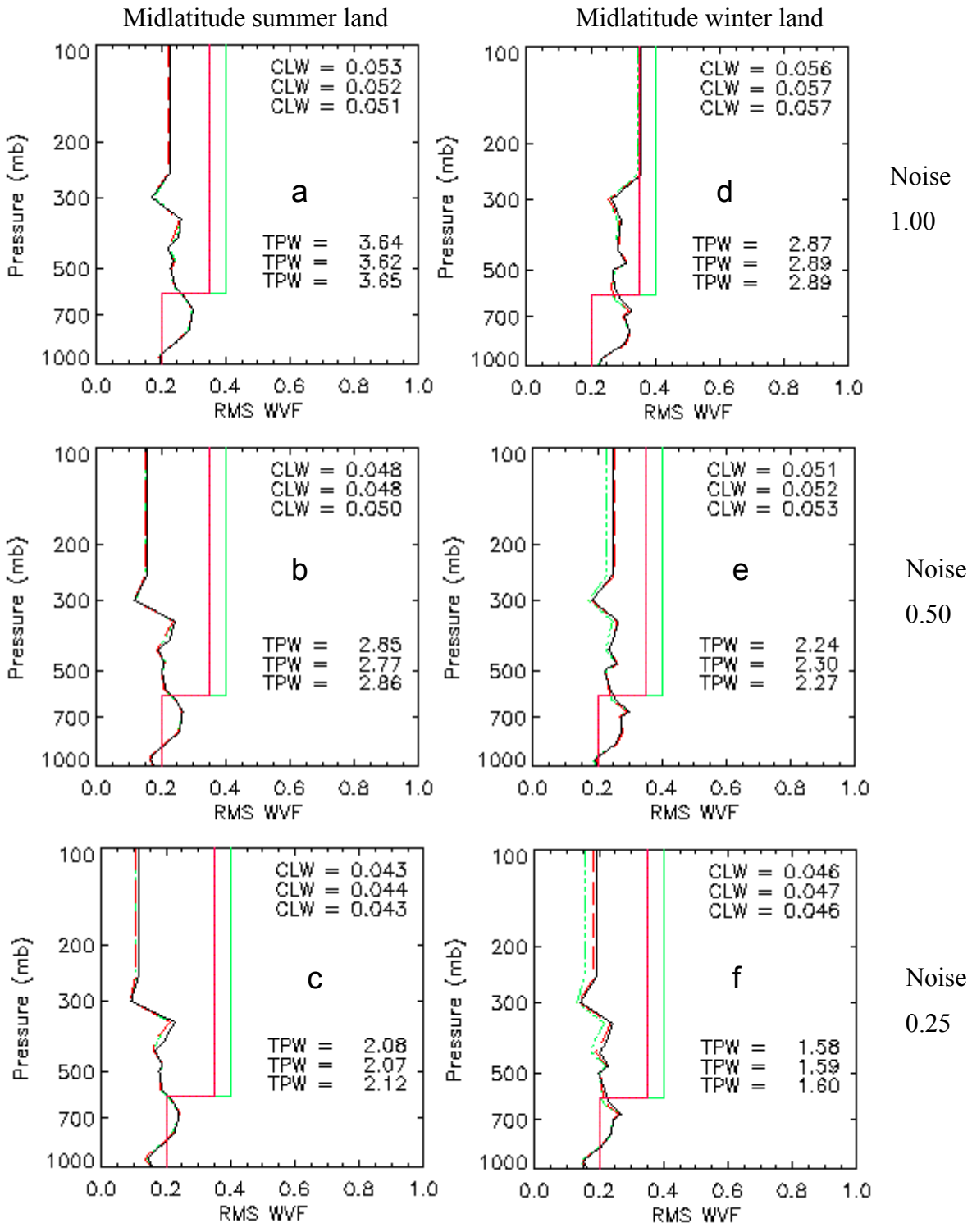


Figure 6-7: Water vapor fractional rms error as a function of pressure, depending on the sensor data noise. The nominal noise was altered by the factor listed at right.

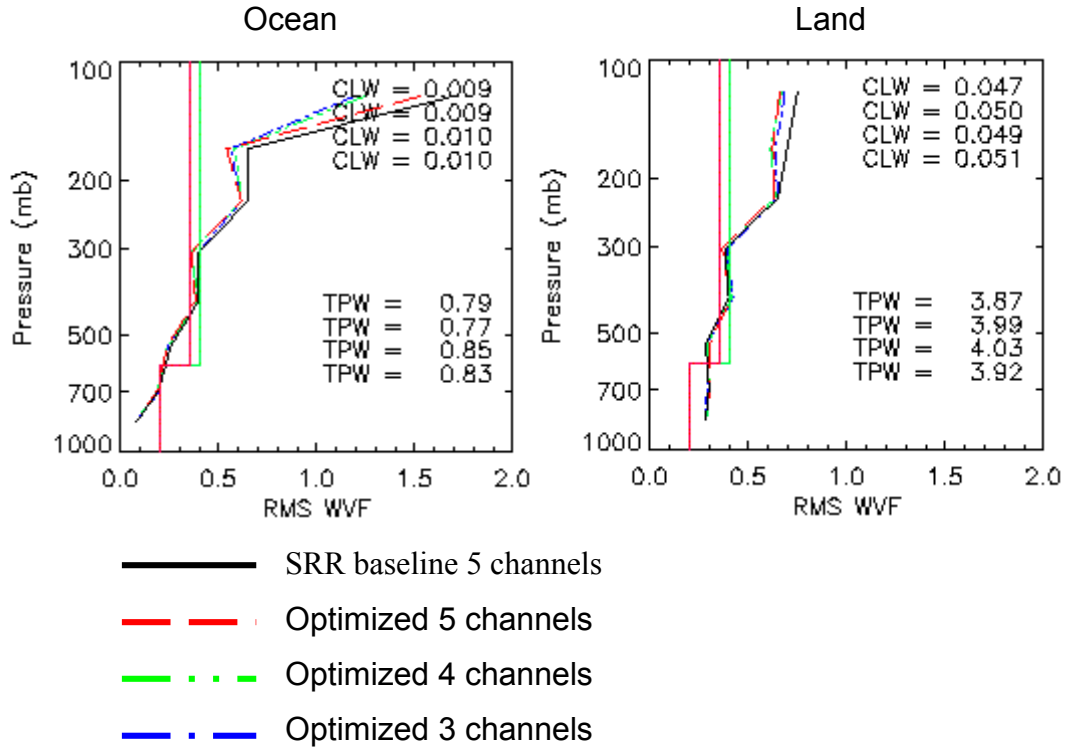


Figure 6-8: Water vapor retrieval performance with the alternative data set. For this figure, the error statistics are for vertically averaged data.

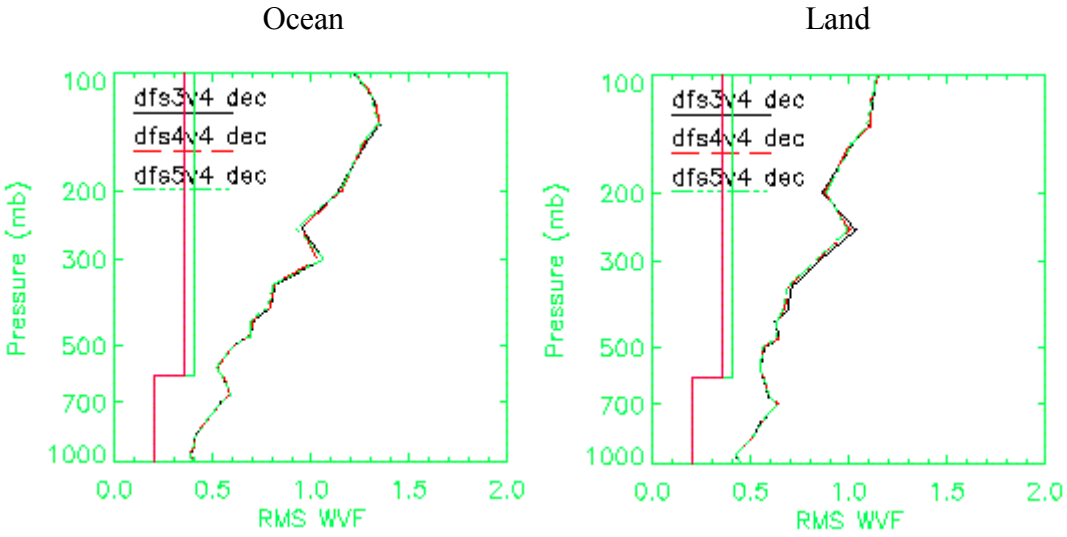


Figure 6-9: Water vapor fractional rms error as a function of pressure, for the perturbed dataset.

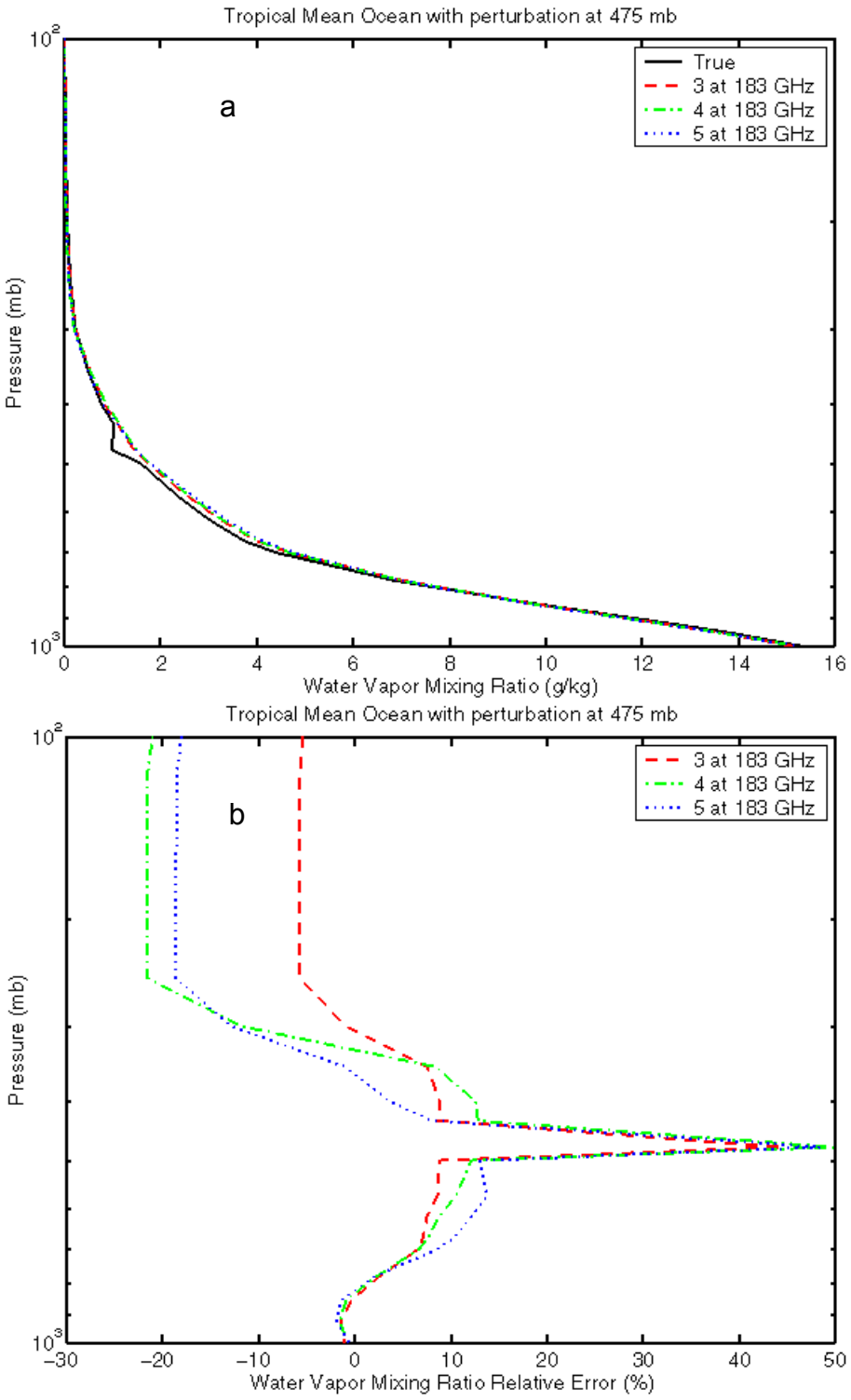


Figure 6-10: Results of retrieval experiments with the optimized 3, 4, and 5-channel sets of 183-GHz channels operating on a perturbed tropical mean profile. Frame a shows the retrieved profiles and the true profile, with the perturbation at 475 mb. Frames b, c, and d show the

retrieval error for mixing ratio, relative to the true mixing ratio for each level, with the frames corresponding to perturbations at 475, 350, and 620 mb, respectively.

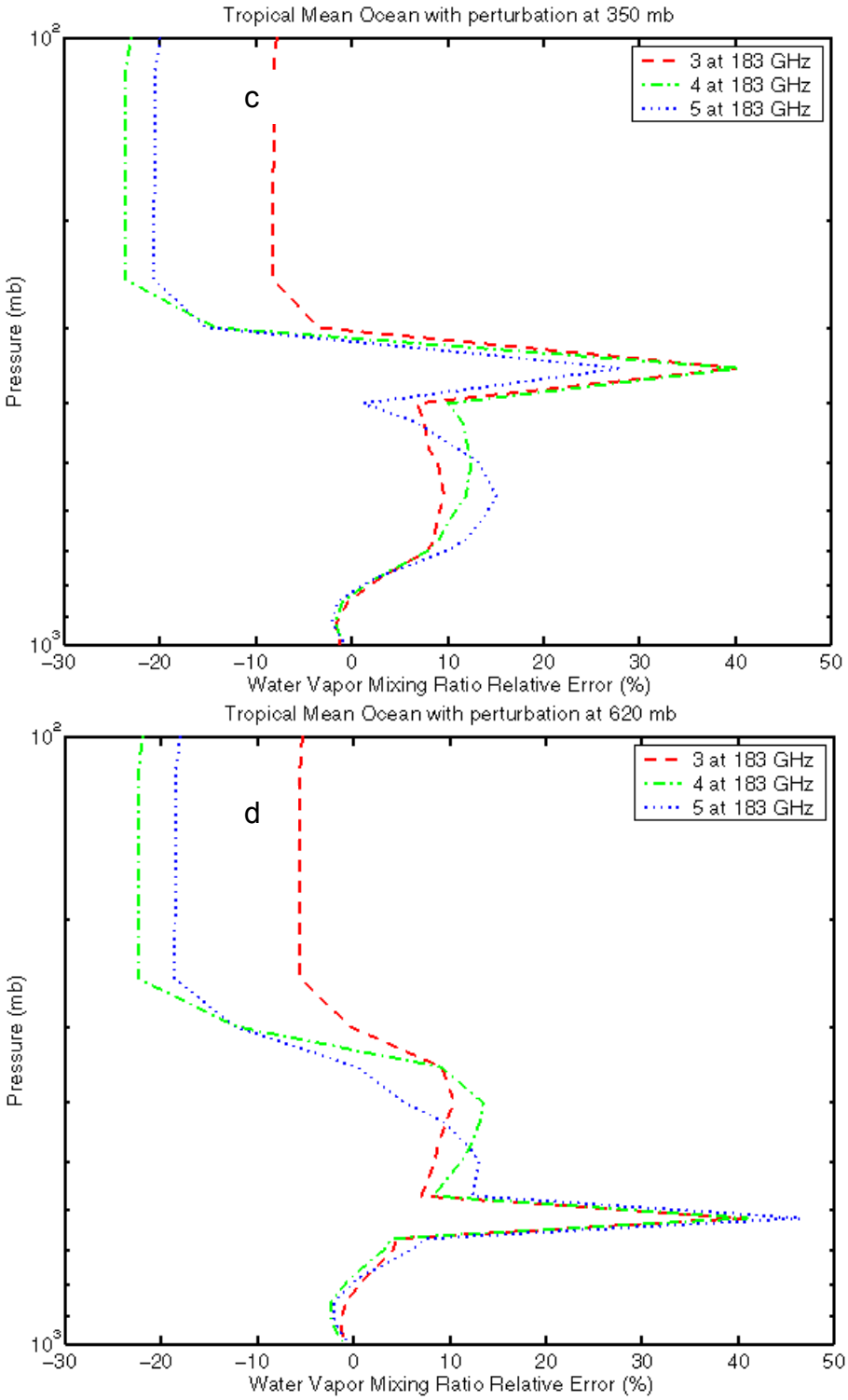


Figure 6-10, continued.

APPENDIX 2: DEFINITION OF 183 GHz CHANNEL SET

APPENDIX 3: ANALYSIS OF 183 GHz CHANNEL SET

From Memo CMIS-AIPT-12a

1 Background

This memo is a followon to an earlier memo on the topic of the 183-GHz channel set, CMIS-AIPT-012v2. That memo evaluated water vapor performance for optimized channel sets having 3, 4, or 5 channels on the 183-GHz line. All of the optimized sets were shown to perform as well or better than the SRR baseline set of 5 channels, which had the same passbands as ATMS. That memo tentatively concluded that the number of 183-GHz channels could be reduced from 5 to 3 without significant degradation of performance. The new material in this memo concerns a revision of the channel set optimization and new experiments regarding performance of channel sets with 3, 4, or 5 channels on the 183-GHz line.

2 Revised Optimization

The channel set optimizations that had been made previously for CMIS had assumed that the radiometer noise figure was constant across the optimization domain, from 0.050 to 10 GHz from the line center. The frequency offset from line center is equivalent to the intermediate frequency (IF) for the double-sideband receiver design. Subsequent to the earlier optimization, BSS provided an estimate of the noise figure as a function of IF (Table 6-1). No data were provided for IF< 0.05 GHz, but it was assumed that NF becomes very large at lower values of IF. The data provided did not include all terms in the noise figure, so we were instructed by BSS to add 0.9 dB to the tabulated values.

The optimization program was modified so that a different NF could be applied to each subband. The modification involved revising the NEDT formula to represent a combination of subbands rather than a single band. When optimizations were run with this new program, the 3, 4, and 5 channel optimal sets each had a passband with an edge at 0.05 GHz IF, despite the runup of NF at low IF. The previous optimization had also resulted in passband edges at 0.05 GHz. However, the new and previous channel sets were not identical with respect to bandwidths. That is, the new and previous channel sets all had the innermost edge of the innermost channel at 0.05 GHz and the outermost edge of the outermost channel at 10 GHz, but the division of the spectrum among the channels between those edges was different in the two versions.

Table 6-1: Noise figure as a function of IF for the 183-GHz recievers.

IF (GHz)	NF (dB)
0.05	12.0
0.1	10.3
0.5	9.5
1.0	9.4
2.0	9.2
4.5	9.2
6.5	9.2
7.5	9.2
8.5	9.2
9.5	9.2
10.5	9.2

The optimizer we use treats passbands as boxcar functions and the results are used to specify the 3-dB bounds of the channels. In the boxcar case, the 3-dB bandwidth, which is a factor in the NEDT budget, is equivalent to the total bandwidth. Channels actually have some finite rolloff rather than a step function at the band edges. A channel with a 3-dB edge at 0.05 GHz would have some significant signal from lower IF. Considering the sharp runup of NF below 0.05 GHz IF, we decided there should be some spectrum between 0.05 GHz and the innermost 3-dB band edge to accommodate the rolloff. Data from BSS indicated that 0.025 GHz would be a sufficient allocation for rolloff, so the optimization was repeated while imposing a lower limit of 0.075 GHz IF. Channel sets optimized this way are listed in Table 6-2.

Table 6-2: Passbands of optimized 183-GHz channels, for sets with 3, 4, or 5 channels.

3 channels		**4 channels**		**5 channels**	
Center (IF) (GHz)	Width (GHz)	Center (IF) (GHz)	Width (GHz)	Center (IF) (GHz)	Width (GHz)
0.7125	1.2750	0.5625	0.9750	0.4625	0.7750
3.1000	3.5000	2.1500	2.2000	1.4500	1.2000
7.7000	4.5000	5.1500	3.8000	3.1500	2.2000
		8.5000	2.9000	6.0000	3.5000
				8.8500	2.2000

3 Performance Evaluations

Water vapor retrieval performance tests were made with a set of 1378 profiles over ocean and another set of 1380 profiles over land. The atmospheric profiles were taken from the NOAA88 dataset and the surface emissivities were from the Prigent October Europe/Africa dataset (see more details about these datasets in the Core Module ATBD appendices). The test profiles contained no cloud. Inclusion of cloud would tend to mask the performance differences among the channel sets, with respect to water vapor profiling. The NEDT values and the noise averaging factors, which depend on footprint sizes, were based on the SFR design.

Results for the ocean and land profile sets are in Figure 6-11. In neither case does a significant difference appear between the results from 3, 4, or 5-channel sets of 183-GHz channels.

Tests with the previous channel set had shown no significant benefit of using any more than 3 channels for the 183-GHz line, as reported in memo CMIS-AIPT-012v2. Some idealized experiments, conducted after the dissemination of that memo, indicated there are some environments where additional channels provide substantial benefits. In particular, water vapor profiles with extraordinarily dry or wet layers appeared to benefit from using a larger number of channels (TIM 9/99 charts).

The previous memo had dealt with the issue of extraordinary profiles to some degree. It reported experiments were we added random variations to individual levels in the profiles. The trouble with that approach was that it added water vapor structure only on the finest vertical scales supportable by the vertical grid. Even a system with many channels would be unable to resolve scales that fine.

We wished to perform experiments with realistic profiles that have a relatively large amount of vertical structure at CMIS-resolvable vertical scales. We did this by computing the eigenvectors and eigenvalues of the covariance matrix of the water vapor profiles, where the profiles were represented as the logarithm of the mixing ratio. The eigenvectors can then be used to transform any given profile into its principal components. The transformation takes a profile with log mixing ratio on 40 levels and converts it to a set of 40 principal components. When the whole

set of profiles is considered, the variance of the first principal component is equal to the first eigenvalue, and likewise for the other principal components.

A profile with an unusually large value for the first principal component is relatively moist or dry overall, depending on the sign. The following several principal components correspond with successively finer scales of vertical structure. Figure 6-12 a) and b) have examples of profiles with relatively large negative and positive values, respectively, of the fourth principal component. In this case, "relatively large" means about 3 times the standard deviation of the principal component (or, the square root of the eigenvalue). The profile in frame a) is relatively moist between 500 and 700 mb, with a sharp drop in mixing ratio below the 700-mb level. The profile in frame b) has the opposite characteristics. Frames c) and d) are for cases with large values of principal component 6, for which the vertical scales are finer than for principal component 4. Principal components of order 11 and higher represent essentially nothing but noise in the profile observational data.

After performing retrievals on the full sets of ocean and land profiles, we filtered the profiles according to the magnitude of their principal components and computed error statistics on the filtered data. Figure 6-13 has results from only the cases where the absolute value of principal component 6 was more than two times the standard deviation of that component. For these cases, the 4 and 5-channel sets performed significantly better than the 3-channel set. There was no significant difference between using 4 or 5 channels.

We examined performance plots for profiles filtered by a range of principal components and found that the distinctions between the 3-channel set and the 4 and 5-channel sets tended to be greatest for profiles with large values of principal components 4 to 8. The lower-order principal components (1 and 2) correspond to broad-scale water vapor structures that can be retrieved with the same skill regardless whether there are 3, 4, or 5 channels in the set. The higher-order principal components (9 and up) correspond to fine-scale water vapor structures for which even sets with 5 or more channels have little skill. Performance is shown in Figure 6-14for the subset of profiles where any of principal components 4 to 8 were greater than their respective standard deviations. This subset makes up 20% of the total data set. There are indications that the NOAA88 profiles are smoother than the true global climatology, so profiles with this degree of structure may be more common in nature than they are in the test dataset. These performance

data indicate an advantage of using 4 channels versus 3, but no significant advantage of using 5 channels versus 4.

For contrast, performance for profiles that were relatively smooth at the resolvable vertical scales is shown in Figure 6-15. This performance is based on the profiles for which none of principal components 3 to 10 had a magnitude greater than two times their respective standard deviations. With the smooth profiles, the performance was independent of the number of channels in the set.

4 Discussion of Results

The errors tended to be larger in the cases with a high degree of structure (Figure 6-14) than for those with a low degree of structure (Figure 6-15). The cases with more structure are more difficult because passive sounders, such as CMIS, do not have great skill at resolving any scales but the broadest ones. While a 4-channel set may resolve more of the variance at a certain scale than a 3-channel set, both channel sets leave a major portion of the variance unresolved. So profiles with more variance at that scale will have more unresolved variance: that is, more error.

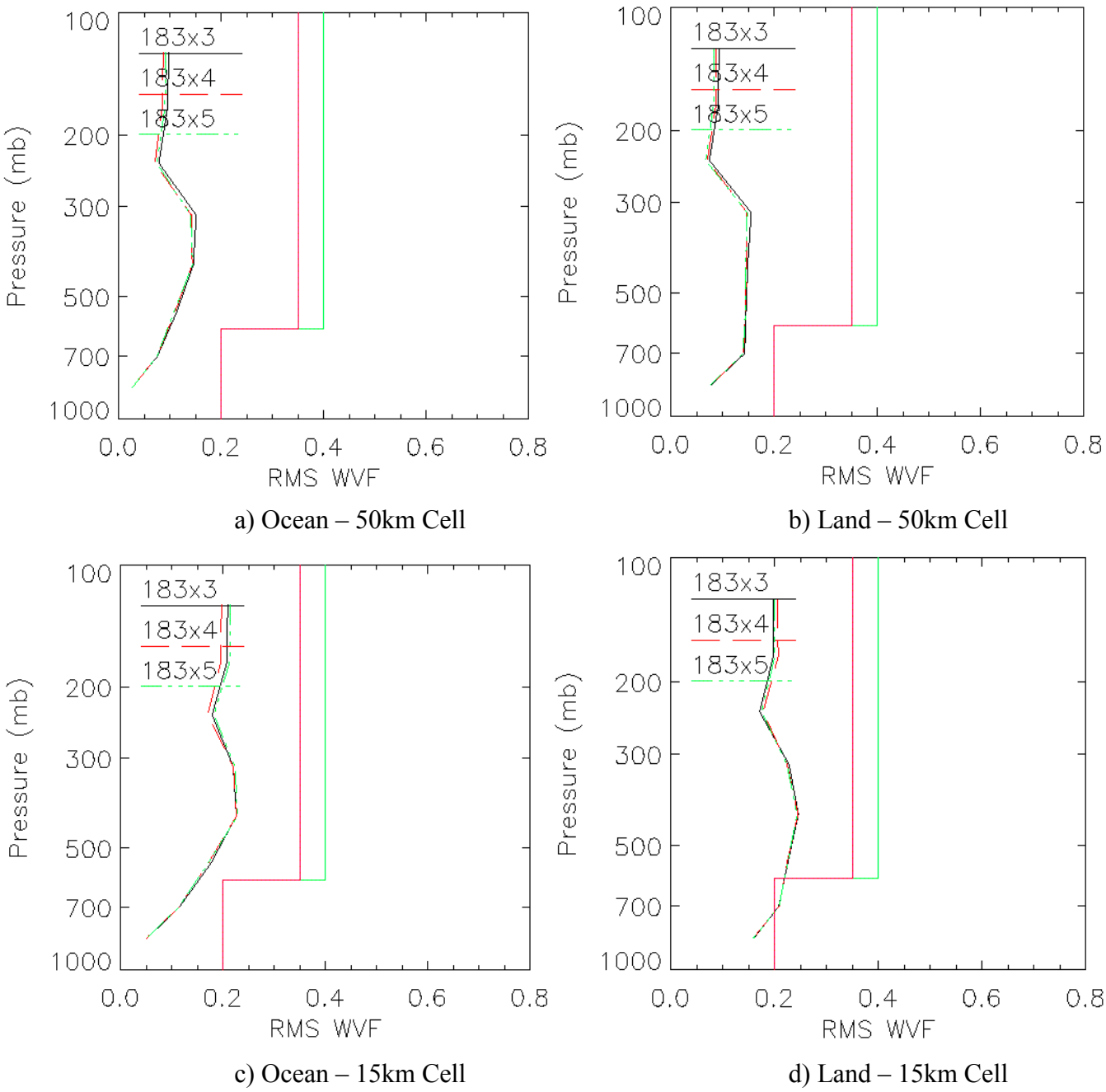
The results in Figure 6-14 show a significant benefit of going from the 3-channel set to the 4-channel set when the retrievals are performed at 50-km cell size, but not when they are performed at 15-km cell size. There are two differences in the controls on retrieval performance at the two scales. The primary difference is that the larger cells are composed of more individual data samples averaged together, so the cell-average NEDT is smaller. A secondary difference is that channels with frequencies below 23 GHz were excluded from the 15-km-cell retrievals because of excessive footprint sizes. We ran tests to show that the latter difference had no significant impact on the relative performances of the 3, 4, and 5-channel sets.

It is well-established (and we have confirmed by analysis; see TIM 9/99 charts) that lower noise corresponds to more independent units of information, when dealing with noisy data such as CMIS brightness temperatures. The additional information that is potentially available in a 4-channel set, compared with a 3-channel set, comes with additional noise. If the noise level is too high, the additional information is lost.

The required horizontal cell size for the Atmospheric Vertical Moisture Profile EDR is 15 km. The results presented here indicate that, if retrievals are performed directly at the 15-km cell size, 3 183-GHz channels are sufficient.

The AER design for EDR algorithms does not call for water vapor retrievals to be performed directly at 15 km. The design calls for performing retrievals first at 40 or 50-km cells and then using those results (interpolated to 15-km) as a constraint on retrievals at 15-km cells. Before commencing a 15-km retrieval, the brightness temperatures computed from the interpolated data are compared with the observed data. If the brightness temperatures differences are small enough, the 15-km retrieval does no work, but just passes on the water vapor profile that was retrieved at 40 or 50 km. The difference will generally be small when the horizontal gradients of water vapor (and other variables) are small. The difference will be larger, and the 15-km algorithm will be executed, when the gradients are relatively large. The 15-km algorithm is executed only when the brightness temperature differences indicate that there is more to be gained, with respect to resolving gradients, than there is to be lost, with respect to increased noise in the retrieval.

With the AER algorithm design, retrieval performance in areas with small horizontal gradients will be similar to the performance shown for 50-km cells. In areas with larger horizontal gradients, performance will approach that shown for 15-km cells. It is not possible at this time to be quantitative about the frequency with which each of these conditions (and all conditions in between) will occur. Quantitative estimates will require further tests with the integrated algorithms applied to test data from weather prediction models that can represent realistic horizontal gradients.



a) Ocean – 50km Cell

b) Land – 50km Cell

c) Ocean – 15km Cell

d) Land – 15km Cell

Figure 6-11: Water vapor fractional error for the 183-GHz optimized channel sets with 3, 4, and 5 channels. Frames a) and b) are for 50-km horizontal cell size and c) and d) are for 15-km horizontal cell size. Frames a) and c) are for ocean background and frames b) and d) are for land background. The profiles were vertically averaged according to the threshold required vertical cell size before computing errors. The red line is the threshold for clear cases.

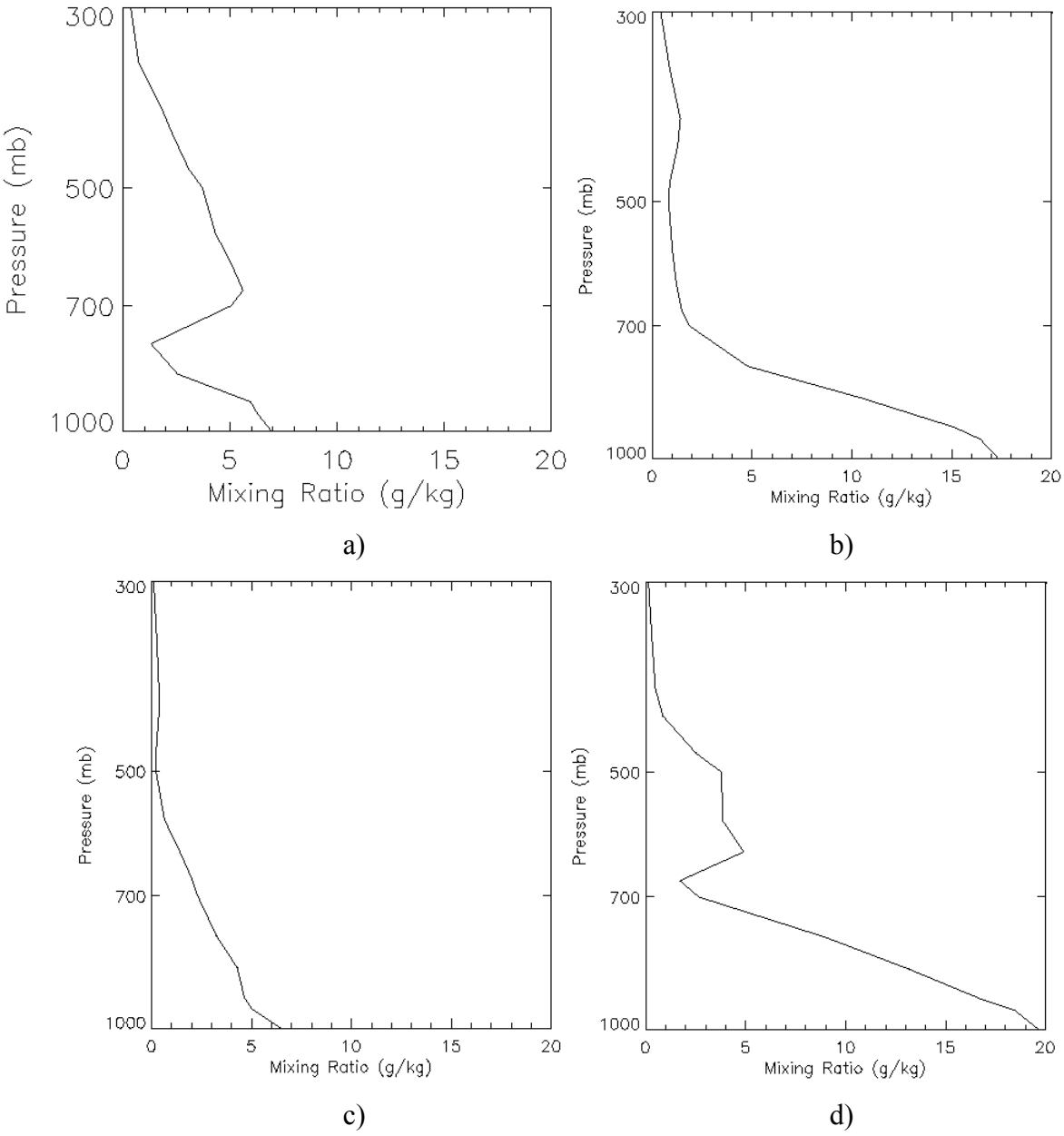


Figure 6-12: Sample profiles of water vapor mixing ratio with exceptionally large values of principal components a) 4, negative, b) 4, positive, c) 6, negative, and d) 7, positive. Note that the choice of sign is arbitrary.

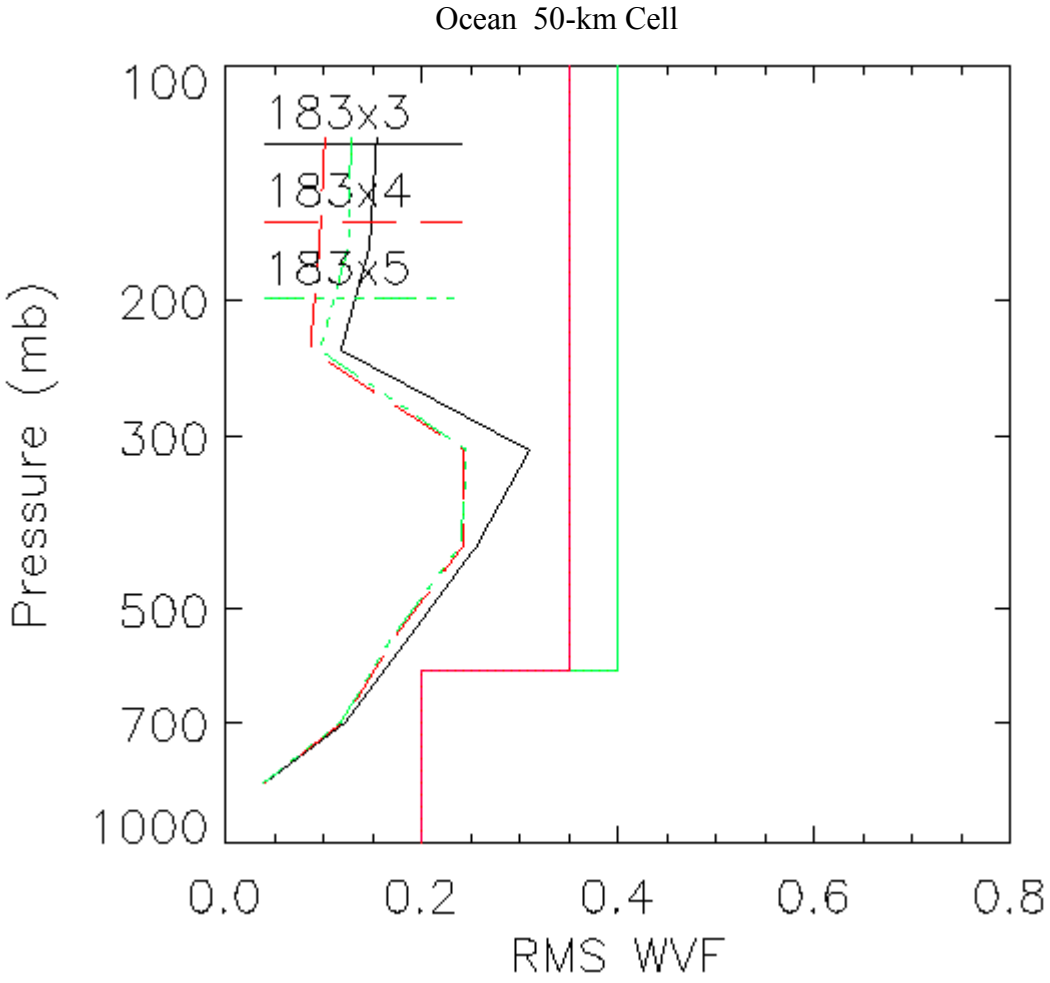
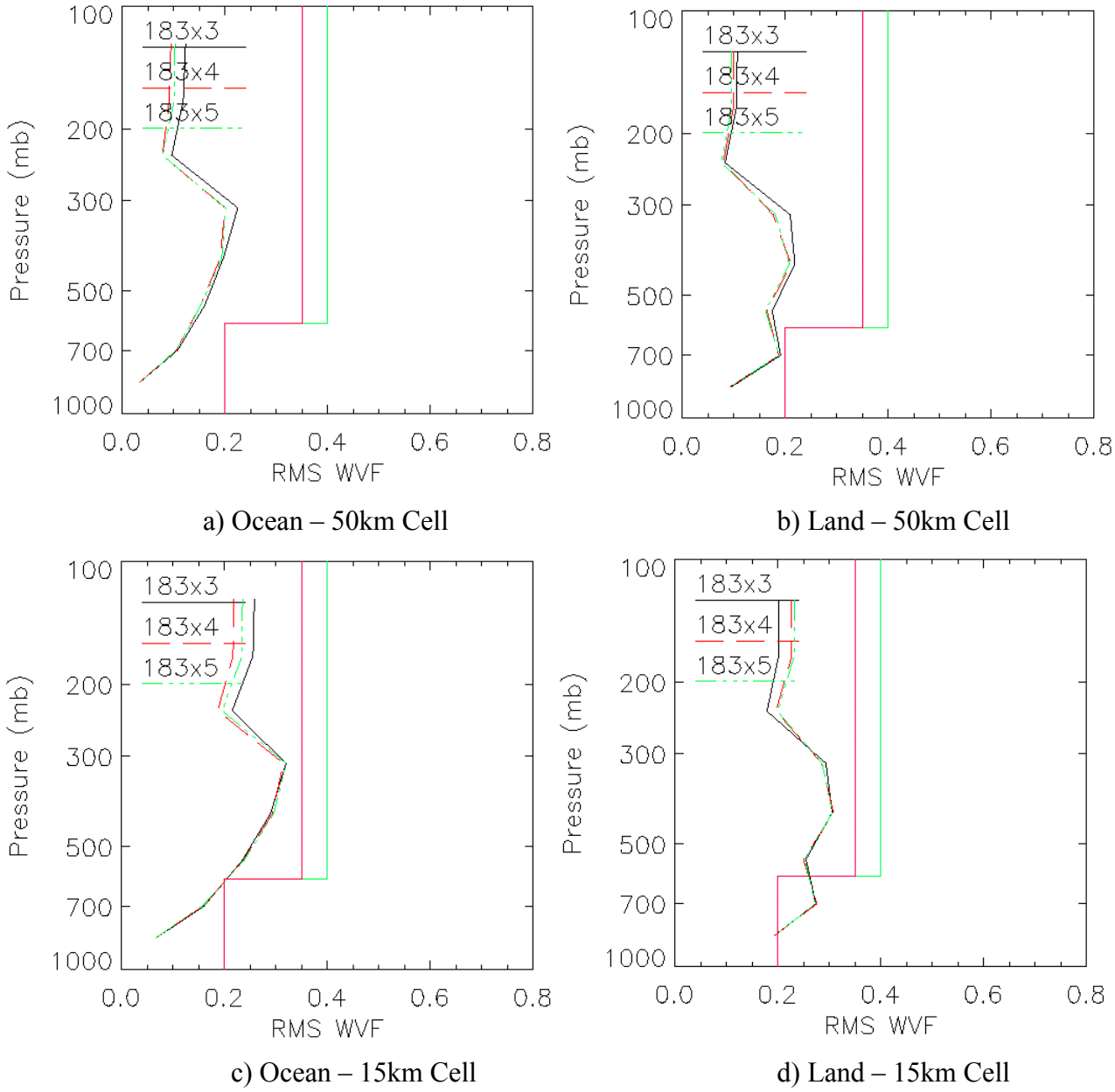


Figure 6-13: Water vapor fractional error for cases where principal component 6 was greater than 2 times the standard deviation (square root of the eigenvalue).



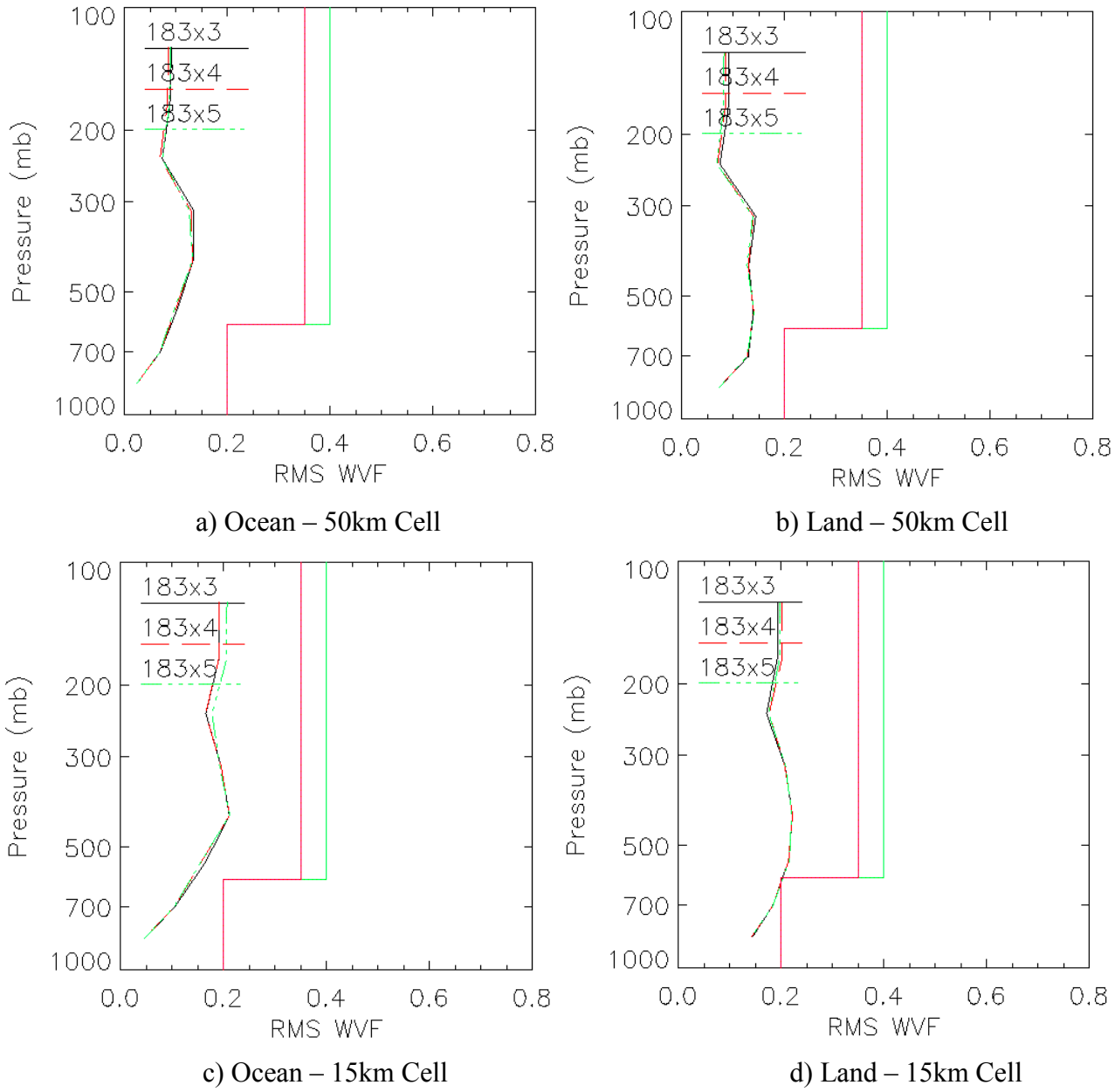
a) Ocean – 50km Cell

b) Land – 50km Cell

c) Ocean – 15km Cell

d) Land – 15km Cell

Figure 6-14: Water vapor fractional error for cases where any of principal components 4 to 8 were greater than 2 times their respective standard deviations.



a) Ocean – 50km Cell

b) Land – 50km Cell

c) Ocean – 15km Cell

d) Land – 15km Cell

Figure 6-15: Water vapor fractional error for cases where none of principal components 3 to 10 were greater than 2 times their respective standard deviations.

APPENDIX 4: SENSITIVITY OF 183 GHz CHANNELS TO WATER VAPOR RETRIEVALS

1 Introduction

The 183 GHz H_2O band is sensitive to moisture profiles in the altitude range from about 800 mbar to 100 mbar depending on the distribution of the moisture in the atmosphere.

The questions that we are trying to answer are the following:

a- How many piece of independent information can be obtained from the 183 GHz H_2O band?

b- What is a best way to perform moisture profile retrievals?

c- What are the pros and cons of using an error covariance matrix in the retrieval?

2 Description of Experiments

To concentrate on the moisture retrieval, all other parameters (surface emissivity, skin and atmosphere temperatures) are presumed known. The vertical and horizontal polarization of 85 GHz channels and vertical polarization of 150 GHz channel are used to sound the lower atmospheric moisture profiles. The 183 GHz H_2O band is divided into 25 double sided channels. Each channel has a spectral bandwidth of 0.6 GHz. The difference between the central frequency and the outer limit of these channels are listed below:

0.3, 0.9, 1.5, 2.1, 2.7, 3.3, 3.9, 4.5, 5.1, 5.7, 6.3, 6.9, 7.5, 8.1, 8.7, 9.3, 9.9, 10.5, 11.1, 11.7, 12.3, 12.9, 13.5, 14.1, 14.7.

3 EOF Analysis of Radiances of 25 Channels in the 183 GHz Band

Radiances of 183 GHz channels were generated using 230 randomly selected ocean profiles with Wilheit ocean emissivity model [Wilheit, xxx]. A flat 0.1 K random noise was added to each channel in the simulation. A principal component analysis was performed on all 230 radiance spectra. Figure 6-16 shows the first 6 principal components or Empirical Orthogonal Functions (EOF). Figure 6-17 shows the associated eigenvalues.

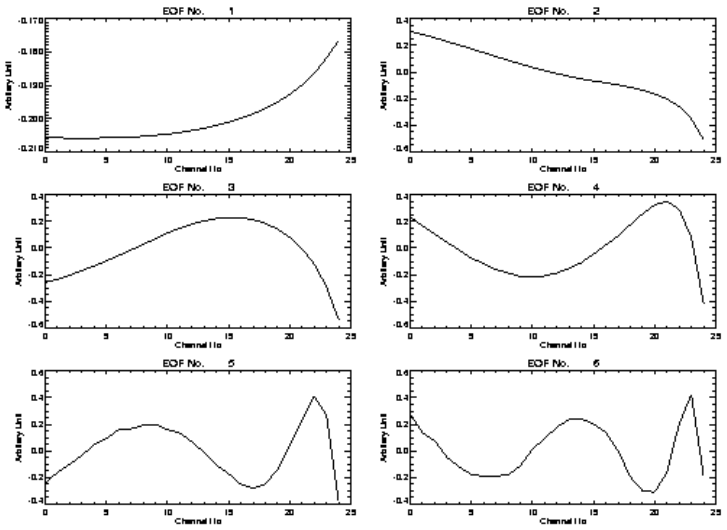


Figure 6-16: The first 6 Principal Components of 183 GHz Radiance Spectra.

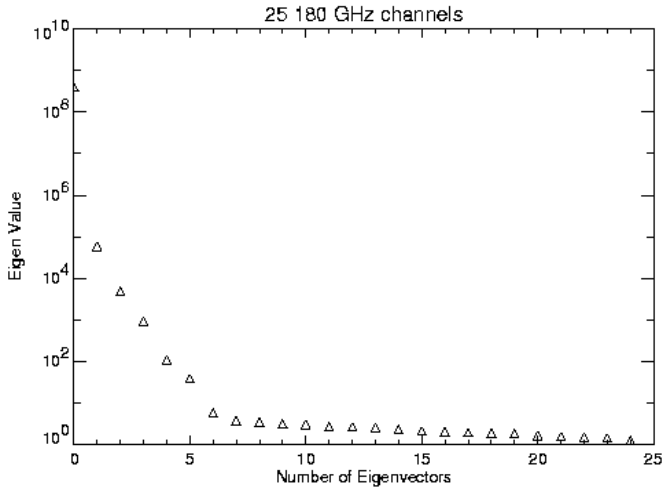


Figure 6-17: Associated Eigenvalues.

The first principal component represents the average of all 230 radiance spectra. The higher order EOFs contribute to variations from the mean spectrum. The contribution of each EOF is proportional to the magnitude of the corresponding eigenvalue. All the EOFs are orthonormal to each other. These figures seem to indicate that there are 6 piece of information that can be distinguished from the noise. But further studies show that the 6th eigenvector may only

associate with a few unique spectra. Using 5 EOFs, one can reconstruct the original spectra within noise level.

Figure 6-18 and Figure 6-19 show that difference between the simulated and reconstructed radiances using 3 and 5 EOFs. The x-axis represents all the spectral points (25 channels multiplied by 230 profiles). It is clear that when using 5 EOFs, the standard deviation of the difference spectra is within the specified random noise (0.1 K). It should be mentioned that if we increase the instrumental noise to 0.2 k, the number of significant EOFs decrease to 4. Studies also show that the above results are unchanged when excluding channels far away from the central frequency (with difference greater than 10 GHz). This indicates that these channels provide no independent information.

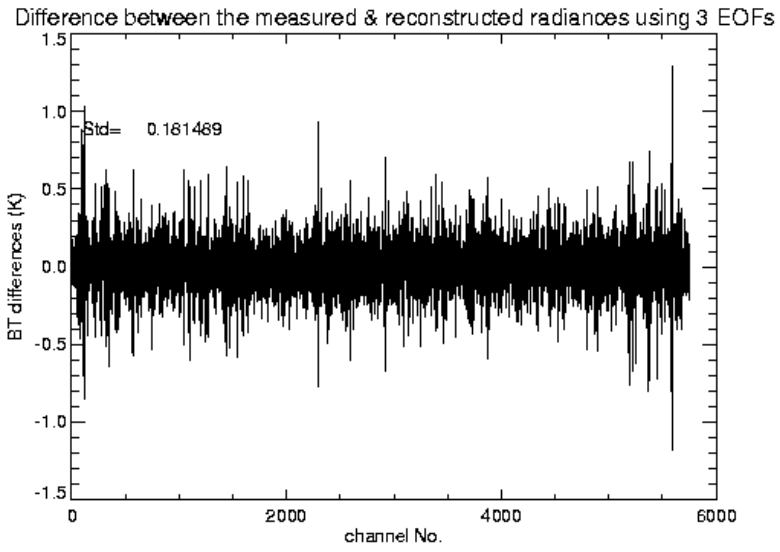


Figure 6-18: Difference between the measured and reconstructed radiances using 3 EOFs.

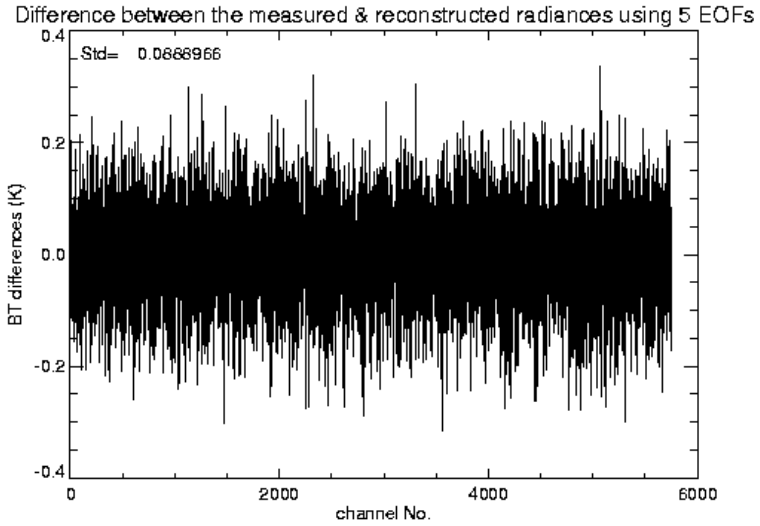


Figure 6-19: Difference between the measured and reconstructed radiances using 5 EOFs.

Based on above study, it is concluded that the 183 GHz band provides 3 to 5 piece of independent piece of information depending on CMIS instrument noise characteristics and the noise reduction factor a certain horizontal cell resolution.

4 Weighting functions of the 183 GHz channels

Weighting functions of the 183 GHz channels are plotted in Figure 6-20. The top curve represents the channel closest to the center of 183 GHz band. As the channel frequency moves away from the band center, the weighting function peaks at lower altitudes. The 3 curves with positive derivatives correspond to 85 GHz (V and H) and 185 GHz channels. It should be mentioned that the altitude distribution of weighting function depends on the moisture profile itself. With moist profiles, the peaks of weighting functions will move higher up in altitude. With drier profiles, the weighting functions at low altitudes have larger values relative to those in the higher altitudes. It is clear that 183 GHz band have little sensitivity for the moisture columns below 900 mbars. Figure 6-20 also shows that, as the channels move away from the band center, the weighting functions tend to peak at about the same height. From retrieval point of view, discarding channels with their frequencies more than 10 GHz away from the band center should have little impact on the performance.

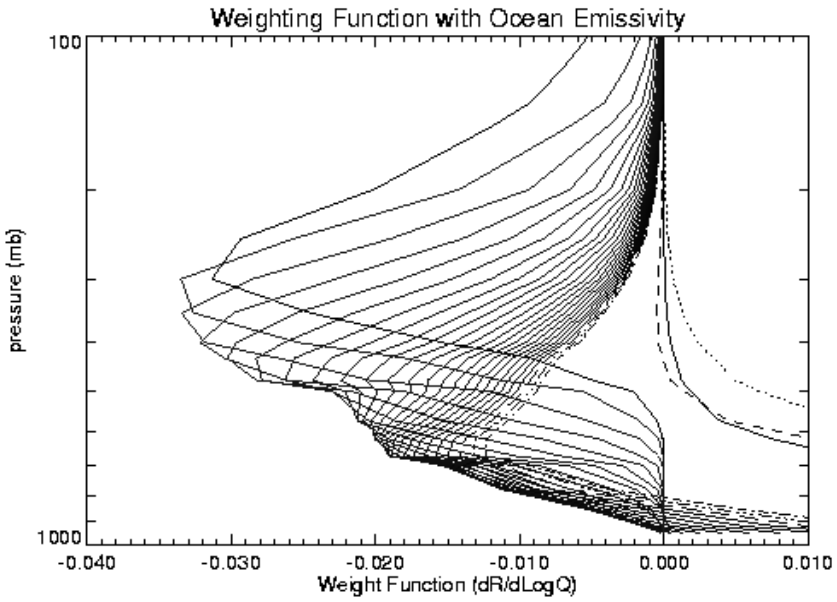


Figure 6-20: A typical plot of weighting function
with a mid-latitude ocean profile and ocean emissivity.

5 Retrieval simulations

The 183 GHz H_2O channels are sub-divided into 3 sets. The first set contains 17 consecutive channels with center frequencies ±0.3, ±0.9, ±1.5, ±2.1, ±2.7, ±3.3, ±3.9, ±4.5, ±5.1, ±5.7, ±6.3, ±6.9, ±7.5, ±8.1, ±8.7, ±9.3, ±9.9 GHz from the band center. The second set contains 6 channels with center frequencies ±0.3, ±1.5, ±2.7, ±3.9, ±5.7, ±8.1 GHz from the band center. The third set contains 3 channels with center frequencies ±0.3, ±3.3, ±6.9 GHz from the band center. In all the retrieval simulations shown below, 2 channels at 85 GHz and one channel at 150 GHz are included to facilitate the moisture retrieval below 800 mb. So the total number of channels for each set are 20, 9 and 6 respectively.

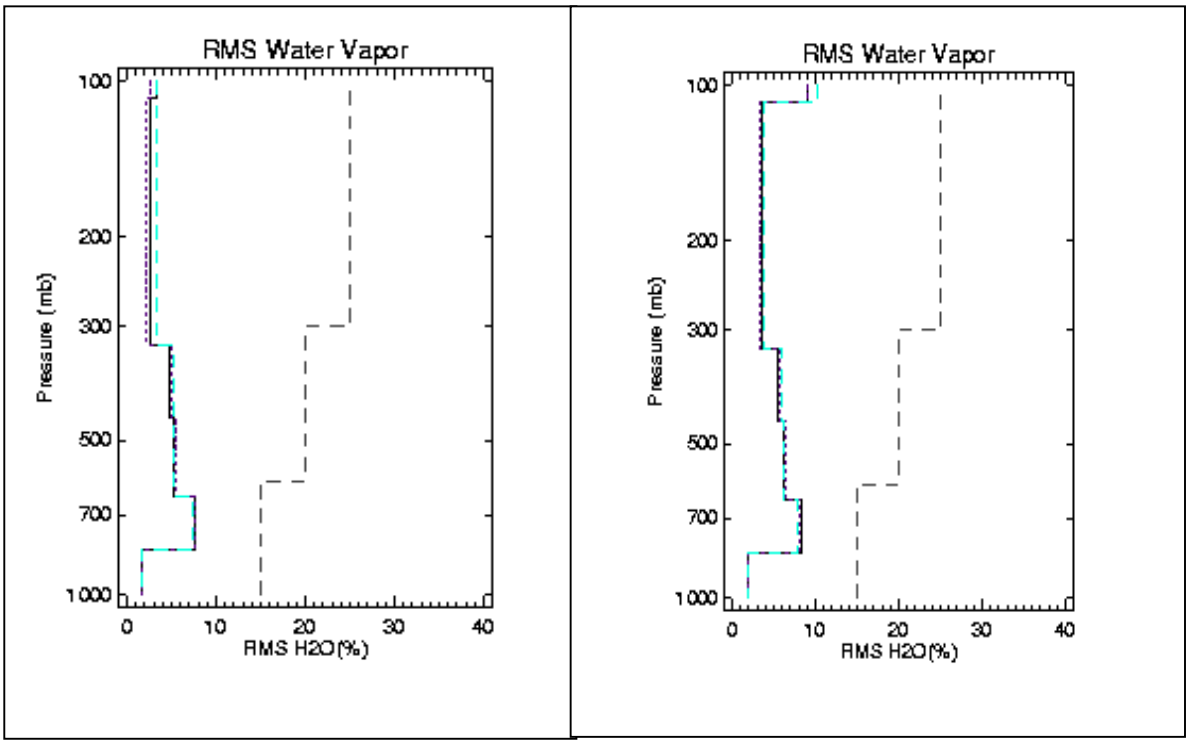
A 0.1 K random noise is added to 85 GHz and 150 GHz channels. The noise added to 183 GHz channels is weighted by the number of channels (nchan) in the band according to:

$$noise = 0.1K * \sqrt{\frac{nchan}{3}} \qquad (13)$$

APPENDIX 4: SENSITIVITY OF 183 GHz CHANNELS TO WATER VAPOR RETRIEVALS

To minimize the impact of uncertainties in temperature and surface emissivity retrievals, these parameters are not retrieved and are set to their true values.

Two methods are used to retrieve moisture profiles. One uses 6 EOF representation and the other uses 6 layers representation. The first method retrieves the projection coefficients of moisture profiles onto the 6 EOFs. The EOFs here are different from the EOFs derived from radiances described in previous section. They are derived from a global climatological moisture profile error covariance matrix. The second method retrieves the layer moisture amount directly in each of the 6 layers. Both are designed to constrain the solution to limited vertical resolution. The purpose of using 2 retrieval methods is to make sure that the retrieval results are consistent between independent retrieval methods. Figure 6-21 a) and b) show that the retrieval RMS using EOF and layer representation respectively.

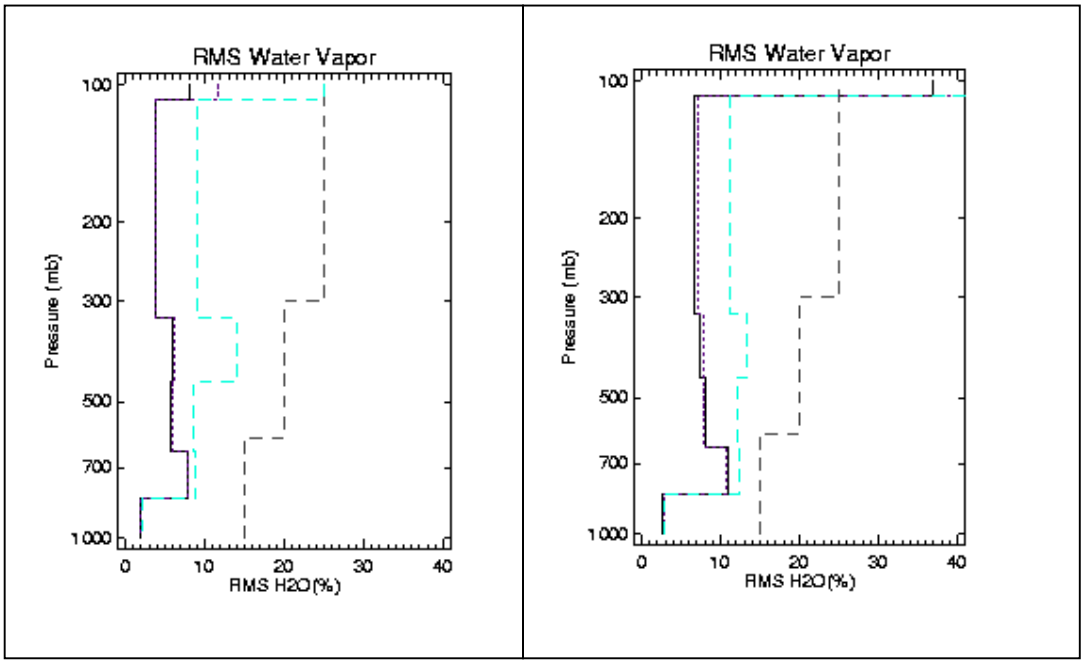


a) b)

Figure 6-21: Retrieval RMS using 6-EOF and 6-layer representations.

The solid lines represent the retrieval RMS using 17-183 GHz channels, the dotted lines represent the RMS using 6-183 GHz channels, and the dashed lines represent the retrieval RMS using 3-183 GHz channels. From these two figures, there is not much difference in the moisture retrieval RMS using different channels set. If we remove the inter-level correlation in the climatology covariance matrix, the retrieval RMS using 3 channels has significantly larger errors (see Figure 6-22 a) and b)).

It's clear from these figures that the moisture retrieval RMS is considerably degraded using 3-183 GHz channels. On the other hand, the retrieval performances using the 6-channel and 17-channel only have small degradation for altitudes above 300 mbar. It is therefore concluded that 3-183 GHz channels do not provide enough information to obtain the best retrieval performance from the radiances.



a) b)

Figure 6-22: The retrieval RMS using 6-EOF and 6-layer representations.
The interlayer correlations in error covariance matrix are removed.

To further determine the optimal number of 183 GHz channels needed for the retrieval, similar studies were done using the channels sets described in technical memo (see Appendix 3). There are 3 "optimized" channel sets with 3, 4 or 5 183-GHz channels. Channels at other frequencies were also used. The NEdT for each channel was calculated according to the bandwidth, the CMIS instrument characteristics, the scene temperature and the noise reduction factor for a given horizontal cell size. The atmospheric temperature profile, the surface skin temperature and emissivity were retrieved in addition to moisture profile. Retrievals were done with 2 horizontal cell resolutions, one with a 50 km cell size and one with a 15 km cell size. Due to the smaller cell size, the noise reduction factors for 15 km cell are about a factor of 5 to 10 smaller than those of 50 km cell.

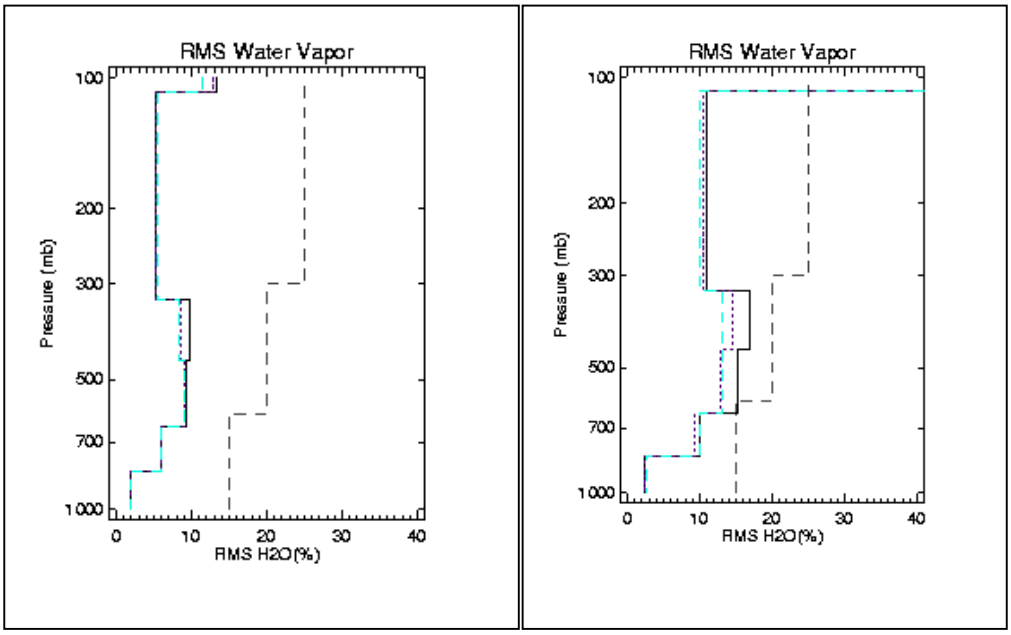
5.1 50 km cell size

Figure 6-23 a) and b) show the moisture retrieval performance using the 6-layer moisture retrieval scheme, with/without interlayer correlations in climatological error covariance matrix. Figure 6-23 b) indicates that when the moisture interlayer correlation is ignored, using 5 183-GHz channels results in better moisture retrieval performance. It's clear that without correlation, the retrieval performance degrades considerably.
There are several reasons for this.

First, the channel noises are higher than the simulations we did before. Therefore, the independent information carried by 183-GHz channels is less. Retrieving 6 independent moisture layer amounts are too much. An appropriate *a priori* information such as climatology background and error covariance matrix is needed.

Secondly, other parameters were retrieved along with 6-layer moisture amounts. It should be mentioned that using 6-EOF method, the retrieval performance is better than the 6-layer method when the interlayer correlation is ignored. This is because these 6 EOFs were generated from the error covariance matrix to start with. A linear combination of the 6 EOFs cannot give a moisture profile with arbitary layer amount in each layer. There is an implicit interlayer correlation built in.

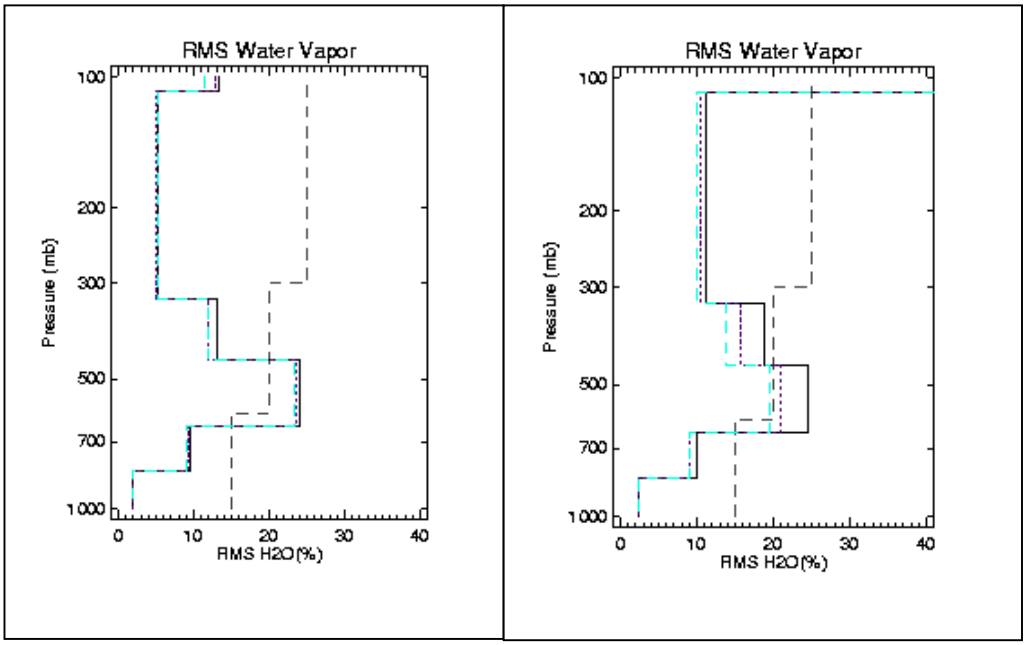
APPENDIX 4: SENSITIVITY OF 183 GHz CHANNELS TO WATER VAPOR RETRIEVALS



a) b)

Figure 6-23: Moisture retrieval performance using the 6-layer moisture retrieval scheme, with (a) and without (b) interlayer correlations in climatological error covariance matrix.

The moisture interlayer correlation in the climatology error covariance matrix apparently helped the retrieval performance. The question is that under abnormal condition, will the correlation included in the climatology error covariance matrix introduce bias in the retrieval? To simulate this, a negative 40% perturbation, in the layer from 430 mb to 570 mb, is applied to the true profiles before the radiance simulation. The retrievals are done with or without the moisture interlayer correlations, and the results are shown in Figure 6-24 a) and b.) Figure 6-24 a) indicates that keeping the interlayer correlation in the error covariance matrix introduces a bias in the perturbed layer. While the retrieval performance, without interlayer correlation and with 4 or 5 channels, improved significantly in the perturbed layer (see Figure 6-24 b)).



a) b)

Figure 6-24: Moisture retrieval performance with (a) and without (b) the interlayer correlation in the error covariance matrix.

5.2 15 km cell size

The channel noise for 15-km cell size is significantly higher than that of 50 km cell size. The retrieval performance for all the parameters is degraded. Similar to the results shown above, when interlayer correlation is included in the error covariance matrix, all 3 183-GHz channel sets give about the same retrieval performance. But when interlayer correlation is not included, the moisture retrieval performance degrades considerably. And there is not much difference between the 3, 4 or 5 183-GHz channel sets.

This implies that one should decrease the number of layers for moisture retrieval when the 183-GHz channel noise increases. Using the 6-EOF method gives better moisture retrieval performance than the 6-layer method, but the result is no better than using the full covariance matrix. Again this is because the 6 EOFs contain implicit interlayer correlation as mentioned before.

The retrieval performance with a 40% perturbation in one layer has also been evaluated. The results show that by removing the interlayer correlation, the 6 EOF method gives better retrieval performance than using a full covariance matrix. The 5 183-GHz channel set gives the best

retrieval performance. On the other hand, the 6-layer method shows no improvement in the perturbed layer. By reducing the number of retrieved layers to 3, we have demonstrated that the retrieval without interlayer correlation gives better results as compared to using a full background error covariance matrix.

REFERENCES

Alishouse, J. C., Snyder, S. A., Vongsathorn, J. and Ferraro, R. R., Determination of Oceanic Total Precipitable water from SSM/I, *IEEE Transactions on Geoscience and Remote Sensing*, **Vol 28, No 5**, pp 811-816, 1990.

Alishouse, J. C., Snider, S. A., Westwater, E. R., Swift, C., Ruf, C., Snyder, S., Vongsathorn, J., and Ferraro, R. R.,Determination of cloud liquid water content using the SSM/I*, IEEE Transactions on Geoscience and Remote Sensing*, **Vol 28, No 5,** pp 817-822, 1990.

Isaacs, R.G., Unified Retrieval Methodology for the DMSP Meteorological Sensors, *RSRM' 87* A*dvances in Remote Sensing Retrieval Methods*, pp 203-214, Deepak, Fleming and Theon Editors, 1987.

Errico, R. M., G. Ohring, J. Derber, and J. Joiner, 2000: NOAA-NASA-DoD workshop on satellite data assimilation. *Bull. Amer. Meteor. Soc*., **81**, 2457-2462.

Grody, N.C., Surface identification using satellite microwave radiometers*, IEEE Trans. Geosci. and Remote Sensing*, **26**, pp 850-859, 1988.

Moncet, J.L. and Isaacs, R.G., Advanced Technical Applications for Satellite Microwave Water Vapor Retrievals, *PL-TR-92-2311, Final Report. Phillips Laboratory, Directorate of Geophysics, AFMC, Hanscom AFB, MA 01731-5000*, 1992.

Moncet, J.-L., Isaacs, R.G., Hegarty, J.D., Unified Retrieval of Atmospheric temperature, Water Substance, and Surface Properties from the Combined DMSP Sensor Suite, *PL-TR-96-2067, 1Final Report. Phillips Laboratory, Directorate of Geophysics, AFMC, Hanscom AFB, MA 01731-5000*, 1996.

Prigent, C., Rossow, W.B., and Matthews, E., Microwave land emissivities estimated from SSM/I observations, *J. Geophys. Res*., **102**, pp 21867-21890, 1997.

Prigent, C., Rossow, W.B., and Matthews, E., Global maps of microwave land surface emissivities: Potential for land surface characterization, *Radio Science*, **33**, pp 745-751, 1998.

Prigent, C. and Rossow, W.B., Retrieval of surface and atmospheric parameters over land from SSM/I: Potential and limitations, *Q. J. R. Meteorol. Soc*., **xxx**, pp xxx-xxx, 1999.

Wentz, F. J., A well calibrated ocean algorithm for SSM/I, *JGR*, **Vol 102**, pp 8703-8718, 1997.

Wentz, F. J., ATBD for the AMSR Ocean Algorithm, *RSS Tech. Report 120296*, 1996.

Wilheit, T.T., The effect of winds on the microwave emission from the ocean's surface at 37 GHz, *J. Geohys. Res*., **84**, pp 4921-4926, 1979.

LIST OF ACRONYMS

ADD	Algorithm Description Document
AER	Atmospheric and Environment Research
AGL	Above Ground Level
ALFA	AER Local Area Forecast Model
AMS	American Meteorological Society
AMSR	Advanced Microwave Scanning Radiometer
AMSU	Advanced Microwave Sounding Unit
APOLLO	AVHRR Processing Scheme Over Cloud Land and Ocean
APS	Algorithm Performance Simulation
ARA	Atmospheric Radiation Analysis
ARD	Algorithm Requirements Document
ARM	Atmospheric Radiation Measurement
ASRR	Algorithm System Requirements Review
ATBD	Algorithm Theoretical Basis Document
ATOVS	Advanced TOVS
AVHRR	Advanced Very High Resolution Radiometer
BT	Brightness Temperature
CEPEX	Central Equatorial Pacific Experiment
CF	Central frequency
CHARTS	Code for High resolution Accelerated Radiative Transfer with Scattering
CKD	Clough, Kneizys and Davies
CLW	Cloud Liquid Water
CMC	Canadian Meteorological Center
CMIS	Conical Microwave Imaging Sounder
COD	Cloud Optical Depth
CTH	Cloud Top Height
CTP	Cloud Top Pressure
DEM	Digital Elevation Model
DMSP	Defense Meteorological Satellite Program
DoD	Department of Defense
DOE	Department of Energy
EDR	Environmental Data Record
EIA	Earth Incidence Angle
EOF	Empirical Orthogonal Function
EOS	Earth Observing System
ESFT	Exponential Sum Fitting Technique
FFT	Fast Fourier Transform
FIRE	First ISCCP Regional Experiment
FOR	Field Of Regard
FOV	Field Of View
GPS	Global Positioning System
GSFC	Goddard Space Flight Center
HIRS	High-resolution Infrared Sounder
HSR	Horizontal Spatial Resolution
IFOR	Instantaneous Field Of Regard
IFOV	Instantaneous Field Of View

IPO	Integrated Program Office
IPT	Integrated Product Team
IST	Ice Surface Temperature
IWVC	Integrated Water Vapor Content
JHU	Johns Hopkins University
JPL	Jet Propulsion Laboratory
LA	Lower Atmosphere
LAT	Latitude
LBL	Line By Line
LBLRTM	Line By Line Radiative Transfer Model
LMD	Laboratoire de Météorologie Dynamique
LON	Longitude
LOS	Line Of Sight
LST	Land Surface Temperature
L-V	Levenberg-Marquardt
LVM	Levenberg-Marquardt
MHS	Microwave Humidity Sounder
ML	Maximum Likelihood
MODIS	Moderate-Resolution Imaging Spectrometer
MODTRAN	Moderate Resolution Transmittance Code
MSU	Microwave Sounding Unit
MW	Microwave
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Center for Environmental Prediction
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NEDN	Noise Equivalent Difference
NESDIS	National Environmental Satellite, Data, and Information Service
NN	Neural Network
NOAA	National Oceanic and Atmospheric Administration
NPM	Numerical Prediction Model
NPOESS	National Polar-orbiting Operational Environmental satellite System
NRL	Naval Research Laboratory
NWP	Numerical Weather Prediction
OD	Optical Depth
OI	Optimal Interpolation
OLS	Operational Linescan System
OSS	Optimal Spectral Sampling
PCA	Principal Component Analysis
PDR	Preliminary Design Review
PSD	Power Spectral Density
POES	Polar Orbiting Environmental Satellite
Psfc	Surface Pressure
PSURF	Surface Pressure
QC	Quality Control
RDR	Raw Data Records

RH	Relative Humidity
RMS	Root Mean Square
RMSE	Root Mean Square Error
RRTM	Rapid Radiative Transfer Model
RT	Radiative Transfer
RTA	Radiative Transfer Algorithm
RTE	Radiative Transfer Equation
RTM	Radiative Transfer Model
S/N	Signal/Noise
SAR	Synthetic Aperture Radar
SCPR	Simultaneous Cloud Parameter Retrieval
SDR	Sensor Data Record
SEIT	System Engineering Integrated Product Team
SFR	System Functional Review
SGI	Silicon Graphics, Inc.
SPS	System Performance Simulation
SRD	Sensor Requirement Document
SRR	System Requirements Review
SSM/I	Special Sensor Microwave/Imager
SSM/T	Special Sensor Microwave/Temperature
SSMIS	Special Sensor Microwave Imager Sounder
SST	Sea Surface Temperature
SVD	Single Value Decomposition
SW	Shortwave
T	Temperature
TBD	To Be Determined (by contractor)
TBR	To Be Resolved (by contractor/government)
TBS	To Be Supplied (by government)
TIGR	Thermodynamic Initial Guess Retrieval
TIM	Technical Interchange Meeting
TOA	Top Of Atmosphere
TOD	Time of Day
TOVS	TIROS-N Operational Vertical Sounder
TRD	Technical Requirements Document
TSKIN	Skin Temperature
UA	Upper Atmosphere
UR	Unified Retrieval
USGS	United States Geological Survey
VIIRS	Visible/Infrared Imager/Radiometer Suite
Vis	Visible
WPTB	Weather Product Test Bed
WV	Water Vapor
WVF	Water Vapor Fraction